From Participation to Motivation: Investigating the Behavioral and Neural Mechanisms of Gaze-Based Social Interactions

Inauguraldissertation

zur

Erlangung des Doktorgrades

der Humanwissenschaftlichen Fakultät

der Universität zu Köln

nach der Promotionsordnung vom 10.05.2010

vorgelegt von

Ulrich Pfeiffer

aus

Schwäbisch Hall, Deutschland

Februar 2013

Diese Dissertation wurde von der Humanwissenschaftlichen Fakultät der

Universität zu Köln im Mai 2013 angenommen

Thesis Advisors

Primary Advisor: Prof. Dr. Gary Bente

University of Cologne

Department of Psychology

Gronewaldstraße 2

50931 Cologne, Germany

Secondary Advisor: Prof. Dr. Dr. Kai Vogeley

University Hospital Cologne

Department of Psychiatry and Psychotherapy

Kerpener Straße 62

50924 Cologne, Germany

Acknowledgements

This thesis would not have been possible without the encouragement and support of colleagues, friends and family. First and foremost, I would like to thank my thesis advisors Kai Vogeley and Gary Bente for their support throughout the last years. I am especially grateful for having been granted the intellectual freedom to pursue my own research interests, as well as the financial freedom to travel around and present the fruit of this work at various exciting conferences. I am also much obliged to Leo Schilbach for his support of my work and for tirelessly sounding the trumpet for a shift of paradigms in social neuroscience. I am further thankful to my colleagues at the Neuroimaging Group in Cologne (Kathi Krämer, Bojana Kuzmanovic, Astrid Gawronski, Ralf Tepest) – and especially my magnificent office mates Alexandra Georgescu and Bert Timmermans – without whom my time in the lab would have been dreadfully boring and my thesis project would have probably gone terribly awry. I am deeply indebted to my friend and colleague Bert Timmermans for proof-reading this thesis and for countless fertile discussions about data, stats, projects, booze and life in general. Admittedly, the latter ones often were slightly less fertile, but all the more fun and enjoyable. I also would like to thank my parents for having always supported me even though they probably had stopped believing long ago that I was ever going to finish this project. Finally, my love and my deepest thankfulness go to my wife Kathleen for her love and for being so patient with me.

Table of Contents

List of Figures

List of Abbreviations

Abbreviation	Meaning
a/pSTS	Anterior/Posterior superior temporal sulcus
ACC	Anterior cingulate cortex
ASD	Autism spectrum disorders
BOLD	Blood-oxygen level response
dlPFC	Dorsolateral prefrontal cortex
EEG	Electroencephalography
FFA	Fusiform face area
fMRI	Functional magnetic resonance imaging
FPAN	Fronto-parietal attention network
IFG	Inferior frontal gyrus
IPL	Inferior parietal lobule
IPS	Intraparietal sulcus
JA	Joint attention
MEG	Magnetoencephalography
MNS	Mirror neuron system
mPFC	Medial prefrontal cortex
NJA	Non-joint attention
PCG	Precentral gyrus
ROI	Region of interest
SA	Shared attention
SBN	Social brain network
TPJ	Temporo-parietal junction
VS	Ventral striatum

List of Studies

Published (thematic ordering):

Study 1 Wilms M*, Schilbach L*, **Pfeiffer UJ**, Bente G, Fink GR, Vogeley K, 2010. It's in your eyes – using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Social Cognitive and Affective Neuroscience*, 5(1), pp. 98 – 107. (*equal contributions)

Contribution U.P. assisted in designing and testing the interactive eye-tracking paradigm, commented on the manuscript, and wrote a section of the manuscript.

Study 2 **Pfeiffer UJ**, Schilbach L, Jording M, Timmermans B, Bente G, Vogeley K, 2012. Eyes on the mind: investigating the influence of gaze dynamics on the perception of others in real-time social interaction. *Frontiers in Psychology*, 3, p. 537.

Contribution U.P. designed the experiments, collected parts of the data, conducted the statistical analyses, created the graphs, prepared a first version of the manuscript, and implemented the comments of the co-authors and reviewers in the final version of the manuscript.

Study 3 **Pfeiffer UJ**, Timmermans B, Bente G, Vogeley K, Schilbach L, 2011. A non-verbal Turing test: differentiating mind from machine in gaze-based social interaction. *PLoS One*, 6(11), e27591.

Contribution U.P. designed the experiments, collected all data, conducted the statistical analyses, created the graphs, prepared a first version of the manuscript, and implemented the comments of the co-authors and reviewers in the final version of the manuscript.

Submitted:

Study 4 **Pfeiffer UJ**, Schilbach L, Timmermans B, Kuzmanovic B, Georgescu AL, Bente G, Vogeley K. Why we interact: On the role of the ventral striatum during real-time social interactions.

Contribution U.P. designed the experiment, performed behavioral pilot studies, collected all data, conducted the statistical analyses, created the graphs, prepared a first version of the manuscript, and implemented the comments of the co-authors in the final version of the manuscript.

1 Introduction

The success of the human species has been attributed to our ability to navigate the social world: we exchange ideas, we collaborate, we compete, we build networks, we learn from one another and teach each other, and we support each other – in brief, we constantly interact and understand one another through social interaction. Surprisingly, empirical research on social interaction has largely focused on tasks in which participants are passive observers and evaluators of static social stimuli while being detached from the actual interaction. This has caused a somewhat ironic situation in which social cognition is examined in the absence of dynamic social interactions (i.e. *offline*) in what has been termed 'isolation paradigms' (Becchio et al., 2010). Although this approach has generated vital insights into human social cognition, recent claims emphasized that the active engagement with others in interaction (i.e. *online*) plays a particular role in understanding other minds (Schilbach et al., *in press*) and might underlie the development of our ability to think about others, rather than the other way round (Reddy and Morris, 2004).

Social cognition and the underlying neural systems have often been related to the disambiguation of other person's behavior (Brown and Brüne, 2012; Frith, 2007). However, there is more to the participation in social interactions than the extraction of information about an interaction partner. This becomes evident by considering how much time we spend interacting with others: We gossip with our colleagues at work and chat with business partners before actually getting down to business. We invite friends for dinner or meet them for a drink at the bar. We engage in small talk with the sales staff at our local grocery store, even though these people are strangers. If we did all that merely to obtain information about our conspecifics or to predict their behavior, social interaction would be quite a tedious and tiresome process. Most readers will agree that this is not the case and that, quite on the contrary, they commonly experience social interaction as enjoyable and satisfying. Realistic studies of social interaction should therefore consider both dynamic *and* motivational aspects of interaction.

One major reason why such studies are scarce is the lack of suitable empirical methods balancing ecological validity and experimental control. This thesis hence has

strived for an understanding of the cognitive mechanisms and neural processes involved in online social interaction by following a methodological as well as an empirical goal. First, an interactive eye-tracking setup was developed which allowed participants to engage in gaze-based interactions with virtual agents in real-time. Gaze was chosen for the operationalization of social interactions due to its key role in non-verbal communication and its easy accessibility (***study 1***). Subsequently, this setup was used to scrutinize the dynamic parameters of social gaze and to evaluate the perception of other person's gaze behavior as mediated by virtual characters (***study 2***). This was a prerequisite for the empirical objective of this thesis: the construction of a realistic, gaze-based interaction paradigm allowing the investigation of the factors contributing to our perception of an interaction as social and the underlying neural mechanisms. This paradigm – a 'non-verbal Turing test' – required participants to distinguish human and computer interaction partners based on a virtual agent's gaze behavior. The experience of social interaction was therein not defined *a priori*, but emerged based on the interaction dynamics. This allowed for an unconstrained assessment of participants' subjective experience of being engaged in a social interaction (***study 3***). These developments finally permitted addressing the neural basis of the motivation to engage in social interactions and their rewarding nature in a combined eye-tracking and fMRI experiment (***study 4***). Prior to the discussion of these studies, an introduction is given into the concepts of offline and online interaction, into social gaze as a means to investigate online interaction, and into the neural mechanisms underlying our understanding of others.

2 Observation vs. Participation: The Distinction between Offline and Online Social Cognition

One of the major research agendas in social and developmental psychology is the investigation of why and how we understand the mental states of others, a capacity commonly referred to as *mentalizing* (Frith and Frith, 2006) or *Theory-of-Mind* (ToM) (Premack and Woodruff, 1978). However, there is a fundamental distinction in the way in which the underlying processes can be studied. In a recent theoretical account, Schilbach and colleagues have used the terms 'offline' and 'online' social cognition to denote this distinction (Schilbach et al., *in press*). Previous research paradigms have largely concentrated on the study of subjects' understanding of other minds without emotional engagement and participation in a real social interaction. In analogy to internet terminology, this socially disconnected mode of cognition has hence been referred to as *offline* social cognition. In contrast, situations in which an individual is emotionally involved with another individual and engaged in a reciprocal social interaction with that individual have been referred to as *online* social cognition.

2.1 Offline Social Cognition

In customary offline tasks, behavioral or physiological measures are applied to inform conclusions about social-cognitive processes while subjects are passive spectators either observing someone performing actions or interacting with another person. This basically divides the study of offline social cognition into two branches – action observation and mentalizing.

2.1.1 The First-Person Perspective: Observation of Actions and the Mirror Neuron System

Whether we observe a single agent acting on an object or an interaction between two agents influences which perspective we take in order to understand the mental states of the agents (Vogeley and Fink, 2003). In the first case, action observation endows us with a first-person

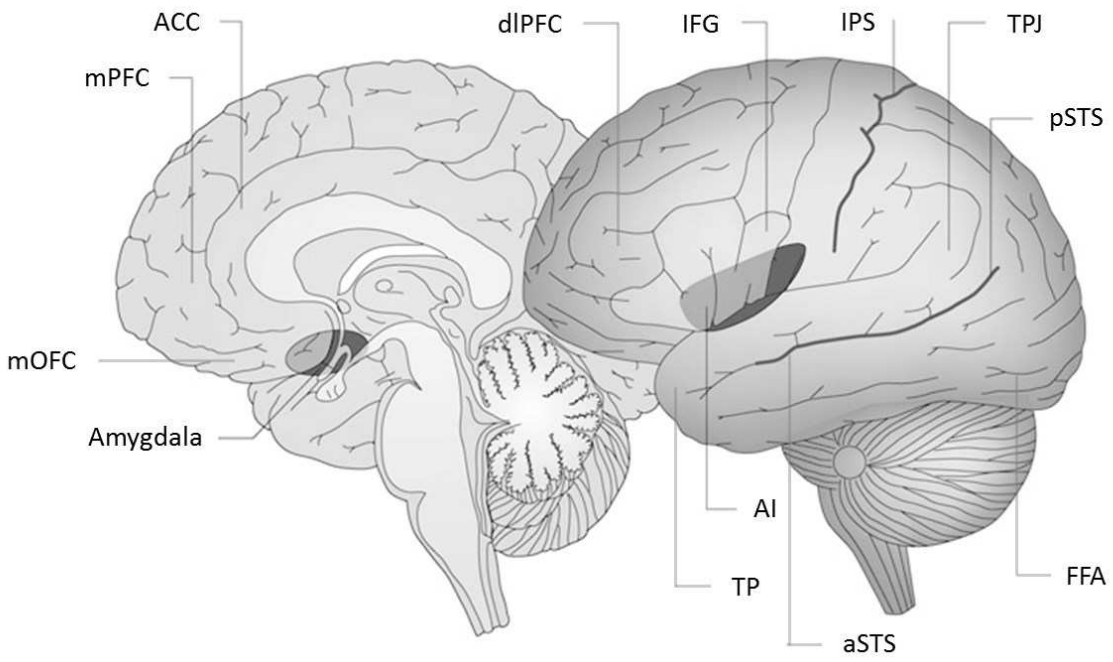

Figure 1. Social brain network and mirror neuron system. Depicted are the key regions involved in social cognition in a medial and lateral view of the brain. See text for details and abbreviations (adapted from Blakemore et al, 2008).

perspective from 'inside' the agent performing the action (Rizzolatti and Sinigaglia, 2010). Neurally, this is supported by the so-called mirror neuron system (MNS), a network consisting of frontal as well as parietal brain areas (Fig. 1). The frontal component of the MNS encompasses the posterior region of the inferior frontal gyrus (IFG) and the inferior region of the precentral gyrus (PCG). The parietal part entails the inferior parietal lobule (IPL) including the inferior parietal sulcus (IPS). Neurons in these areas are active both during the execution of a given action and the observation of another agent performing this action (Di Pellegrino et al., 1992) and hence are believed to coalesce the perception and execution of movements (Rizzolatti and Craighero, 2004). Importantly, when these movements are goal-directed, they are commonly referred to as *motor acts*. In non-human primates, mirror neurons do not simply encode the physical or spatial properties but rather the goal of these movements, i.e. the motor acts. For example, when the goal of a movement was to grasp an object, the neuronal output of neurons of the MNS was not significantly affected by the exact way the fingers were moved or by the tool which was used to grasp the object (Rochat et al., 2010; Umiltà et al., 2008). Although in humans the MNS seems to encode movements *per se* as well as motor acts (Rizzolatti et al., 1996; Gazzola et al., 2007;

Liu et al., 2007), there is common acceptance that the MNS supports the understanding of the goals and intentions underlying the motor behavior of another agent via low-level processes in areas of the motor and parietal cortices (Rizzolatti and Sinigaglia, 2010). Considering that movement "is the only means we have to interact with both the world and other people" (Wolpert et al., 2001 p. 487), it can be assumed that the evolution of a system dedicated to the understanding of actions and their underlying intentions has been of great significance for the phylogenetic development of the human brain.

Nevertheless, there is evidence that the MNS might only support the comprehension of familiar actions taking place in stereotypic contexts. A study by Brass et al (2007) examined the neural correlates of action observation using stimuli with unfamiliar *transitive* actions (i.e. complex movements involving objects) in non-stereotypic contexts with varying plausibility. For example, they showed images of a person pushing a light switch with their knee while their hands were free, implausibly occupied, or plausibly occupied. Thereby, the difficulty of understanding an action was manipulated. Results indicated that understanding the intention of novel actions in implausible contexts requires a different mechanism than a simple mapping of these actions onto motor representations in the MNS. When the plausibility of an action was low, there was significantly increased activation of the superior temporal sulcus (STS) and the medial prefrontal cortex (mPFC). This was corroborated by findings demonstrating that the observation of simple, *intransitive* actions (e.g. minimal finger movements) only recruited the MNS when the context was plausible, but correlated with activation of the temporo-parietal junction (TPJ) when the same actions were observed in an implausible context (Liepelt et al., 2008). The STS, mPFC, and TPJ are key constituents of the so-called 'social brain network' (» Chapter 2.1.2) and recruited whenever mental states need to be actively inferred across a range of social-cognitive tasks (Frith and Frith, 2006). Offline social cognition therefore does not only rely on a first-person grasp of motor intentions, but also on a context-sensitive inferential system (De Lange et al., 2008).

2.1.2 The Third-Person Perspective: Mentalizing and the Social Brain Network

In contrast to the observation of actions exerted by a single individual, the observation of an interaction between two individuals has been suggested to endow us with a third-person perspective on the mental states that drive and emerge during interactions (Frith and Frith,

2006). Mentalizing about others from a third-person perspective encompasses the active inference of their desires, goals, intentions, or motivations regarding objects and persons in the environment. The empirical study of this branch of offline social cognition does not rely on the presentation of pictures or movies of motor acts, but involves multiple forms of verbal and non-verbal stimuli depicting static or dynamic social scenes (cf. Carrington and Bailey, 2009). The prototypical structure of a mentalizing task is pertinently illustrated by false-belief tasks, which require the understanding that what one knows or believes might be different from what another person might know or believe (Wimmer and Perner, 1983). In the famous Sally-Anne-Task (Baron-Cohen et al., 1985), two dolls named Sally and Anne are introduced to children. In a brief comic strip, Sally hides a marble in a basket. After Sally has left the room, Anne takes the marble out of the basket and hides it in a box. Children are then asked whether Sally will search for the marble in the basket or in the box. In order to pass the task, a child needs to abstract from its own knowledge, adopt the perspective of Sally and indicate that she will look for the marble in the basket. This is taken as a proof of intact and fully developed mentalizing abilities. Children younger than four years indicate that Sally will look in Anne's box. This demonstrates that they are unable to understand that another person has beliefs about the world differing from their own and suggests a distinct developmental time course of mentalizing capacities.

2.1.2.1 The Role of the mPFC in Mental State Attribution

Earlier work in social neuroscience has borrowed from the structure of false-belief tasks to construct text vignettes or comic strips depicting social interactions which require the attribution of mental states by active inference (Fletcher et al., 1995; Gallagher et al., 2000; Gobbini et al., 2007; Happé et al., 1996; Vogeley et al., 2001). Although different prefrontal and temporal cortical areas showed enhanced activity during mentalizing, only the mPFC was consistently recruited in *all* of these studies. It has therefore been considered as the hub of the so-called social brain network (SBN) which subserves our capacity to think about the thoughts of others (Amodio and Frith, 2006). Recently, a more complete picture of its functional segregation has emerged. The dorsal mPFC seems to be involved in more explicit and cognitively more demanding aspects of mentalizing. These include mental state inferences about people dissimilar to us (Mitchell et al., 2006), anchoring mechanisms and adjustment processes in cases where others' mental states differ from our own (Tamir and

Mitchell, 2010), and the encoding of communicative intentions that need to be interpreted in a social context as compared to personal intentions (Walter et al., 2004). On the contrary, the ventral mPFC is engaged in implicit aspects of mentalizing, such as reasoning about people we perceive as similar to ourselves as well as the comprehension of the affective states of others (Shamay-Tsoory and Aharon-Peretz, 2007). However, although these results indicate a prominent role of the mPFC in mentalizing, other studies have cast doubts on the specifity of mPFC activation.

2.1.2.2 The Role of the TPJ in Deducing the Content of Mental States

Studies by Saxe and colleagues (for an overview see Saxe, 2006) emphasized the relevance of the TPJ in differentiating the actual *content* of other people's mental states. In a study by Saxe and Kanwisher (2003), more elaborate categories of stimuli were used to isolate mentalizing from other thought processes regarding human behaviors. These categories involved mentalizing, human actions, physical inferences, and physical descriptions. In the central experiment of the study, stories about people which required thinking about either the desires (mentalizing) or the physical properties (non-mentalizing) of the protagonists were compared. Results showed that TPJ activity was increased only when the mental states, but not the physical properties, of a person had to be taken into account. In a later study, the activity of the right TPJ was selectively modulated by thinking about the mental states, but not by thinking about the social backgrounds of people (Saxe and Wexler, 2005). Other regions previously implicated in mentalizing (i.e. the left TPJ or the mPFC) did not display differential activation, thereby suggesting that only the right TPJ is truly selective for mentalizing in the strict sense of understanding the discrete intentions of another person. Interestingly, the strong functional specification of the right TPJ for understanding mental states in adulthood develops rather late in childhood (i.e. between 6 and 11 years). In contrast, the mPFC displays an early sensitivity for reasoning about mental states but shows less developmental change and is less specifically activated during mentalizing in adulthood (Saxe et al., 2009). In the authors' words, this implies that "the basic cognitive signatures of domain-specifity may be in place long before the brain systems underlying these processes have reached an adult-like state" (Saxe et al., 2009, p.1207). In a meta-analysis of mentalizing studies, Van Overwalle and Baetens (2009) suggested that the TPJ is mainly engaged in the inference of transient goals, beliefs, or desires, while the mPFC is recruited

by more spontaneous trait inferences or additional deliberative reasoning. Nonetheless, there is not yet a unifying account of how the interplay of these regions enables mentalizing.

2.1.2.3 Beyond and Below Mentalizing: Other Components of the Social Brain Network

There are several other brain areas associated with the SBN that have been found active during mentalizing tasks. However, their activation was less consistent across these tasks, but was associated with other aspects of social cognition rather than the inference of mental states (i.e. mentalizing *proper*). For instance, the temporal poles are often activated during classical mentalizing tasks (Frith and Frith, 2003), but do not appear to be specifically related to mentalizing. More likely, they are important for the generation of a larger semantic context for the mentalized material by providing social 'scripts' – stereotypical sequences of events which are generated by experience (e.g. 'going to a restaurant') and support social categorization (Schank and Abelson, 1977). Among the SBN regions are also the fusiform face area (FFA), which is implicated in the orienting to and processing of faces (Schultz et al., 2003), and the anterior cingulate cortex (ACC), which plays a fundamental role in the resolution of conflicting social information (Zaki et al., 2010) as well as the monitoring of incongruent self- and other-perspectives and inhibition of undesired mental states of others (Hartwright et al., 2012). The dorsolateral prefrontal cortex (dlPFC) is involved whenever top-down modulation is needed to overcome automatic social-cognitive processes such as racial prejudices (Cunningham et al., 2004) or negative responses to unfair treatment (Kirk et al., 2011). Although they are numbered among the SBN regions, the ACC and the dlPFC exert more domain-general functions not exclusively related to social cognition. This is equally true for the amygdala, which plays a more general role in encoding behaviorally relevant aspects of sensory stimuli (Adolphs, 2010). However, it bears specific relevance in social cognition by encoding socially salient features, such as emotional facial expressions (Fitzgerald et al., 2006) and especially fearful faces which might signal threat (Adolphs et al., 1994; Morris et al., 1996). Finally, early reports indicated that the posterior region of the superior temporal sulcus (pSTS) is involved in the detection and analysis of all aspects of biological motion, such as movements of hand, mouth, eyes and body (Allison et al., 2000). A very recent study suggested that the pSTS is activated by virtually all classes of socially salient stimuli, such as faces, bodies, biological motion, goal-directed actions, emotions, pain, and observed interactions (Lahnakoski et al.,

2012). Despite its seemingly universal importance in offline social cognition, one of the main functions of the STS is the processing of gaze direction (Nummenmaa and Calder, 2009). This functional aspect is highly relevant for face-to-face interactions and is discussed in depth in the chapter on social gaze (» Chapter 3).

2.2 Online Social Cognition

Most readers will not dispute that the observation of actions and interactions contributes to our understanding of others. However, they will probably also agree that everyday social life focuses on the engagement in social interactions with others face-to-face and in real-time (i.e. online) rather than on the inference of the mental states of others from an observer's perspective. An early description of a phenomenological difference between the passive observation of an interaction and the active participation – that is, the singularity of the relation between 'me and you' (i.e. a dyadic interaction) – can be found in the work of the religious philosopher Martin Buber. Although not exactly a social scientist, in his book *I and Thou* (written in 1919 under the German title "Ich und Du" and first translated to English in 1937) he noted that every object or person can be regarded as an 'It' or a 'Thou' and that the fundamental qualities of a relation differ depending on whether it is an 'I – It' or an 'I – Thou' relation (Buber, 2004). In the first case, the other as 'It' is experienced as remote and consisting of a variety of different features that can be observed and pondered upon. In the latter case, the other as 'Thou' is experienced as a whole entity embedded in the same reality as 'I'. This is perceived as the most existential type of relation between two individuals, i.e. one that is situated, embodied and enactive. In Buber's own ornate words, "[t]he *Thou* meets me. But I step in direct relation with it. Hence the relation means choosing and being chosen (…). The primary word *I – Thou* can be spoken only with the whole being. Concentration and fusion into the whole being can never take place through my agency, nor can it ever take place without me. I become through my relation to the *Thou*; as I become I, *I* say Thou. All living is meeting" (Buber, 2004, p. 17).

Over the past few years, a paradigm shift in the study of social cognition has begun to manifest itself, because growing numbers of researchers subscribe to the notion that "it is in engagement with other people rather than in thought that people normally and fundamentally know other people" (Reddy and Morris, 2004, p. 657). Philosophers,

psychologists and neuroscientists alike have advanced the idea that it is not sufficient to study single minds in 'isolation paradigms' (Becchio et al., 2010), but that studies on social cognition should focus on situations in which individuals participate in some form of *online* social interaction (De Jaegher and Di Paolo, 2007; De Jaegher et al., 2010; Froese and Fuchs, 2012; Hobson, 1991; Marsh et al., 2009; Reddy and Morris, 2004; Riley et al., 2011; Schilbach et al., *in press*).

2.2.1 The Second-Person Approach to Other Minds

In the most comprehensive research agenda for the study of online social cognition so far, Schilbach and colleagues put forward a 'second-person approach' to social neuroscience (Schilbach et al., *in press*). The term 'second-person' was coined to clearly demarcate the approach from other accounts presupposing a first- or a third-person perspective (» chapters 2.1.1 and 2.1.2). It emphasizes the fact that online social cognition requires a person to be able to observe the effects of her behavior on the other and to be addressed reciprocally by the other's reactions (i.e. from a second-person perspective as "you"). More specifically, two central components are named in the second-person approach of social interaction – *emotional engagement* and *participation in interaction*: First of all, emotional responses are argued to provide experiential access to the bodily and affective state of the interaction partner due to a tight coupling between the bodily expression of an affective state and the experience of this state referred to as emotional embodiment (Niedenthal, 2007). During social interaction this emotional embodiment can resonate across individuals. Interaction partners unconsciously imitate the emotional expressions of each other and consequently indicate experiencing the associated affective state (Rapson et al., 1994). Secondly, the participation in social interactions is considered a key to the understanding of other minds. In this framework, a social interaction is described as any relation between two (or more) agents that is characterized by a reciprocal exchange of socially salient information via verbal and non-verbal cue systems. There are virtually no limits to the complexity of social interactions. They can range from simple, sequential turn-taking games to non-linear and dynamic procedures such as, for example, political negotiations or a jam session by a group of musicians, who need to keep track of minute changes in rhythm and pitch in order not to end up in cacophony. This illustrates another aspect of online social interaction: It is a truly dynamic process in which the participating agents have to update their intentions and

motivations 'on the fly' during the interaction in order to maintain the interaction. A final feature of social interactions is that the interacting agents usually are either aware of a shared reality (Hardin and Higgins, 1996) and therefore of common resources as well as distinct action possibilities of the other (Sebanz et al., 2006), or strife to gain this awareness by establishing common ground by means of the interaction procedure (Clark, 1996). Beyond the establishment of common ground, it has been argued that some aspects of an understanding of the other reside in the interaction dynamics *per se* rather than on a mere inferential online interpretation of action-reaction feedback loops (De Jaegher et al., 2010). This means that not all social knowledge resides within the minds of individuals but that social cognition is an emergent property of the interaction, thereby assigning an epistemic value to the interaction process itself.

2.2.2 Experimental Evidence for a Second-Person Grasp of Other Minds

The hypothesis that we gain access to others' minds via social interaction is intuitively intriguing. However, every hypothesis about cognition eventually must be supported by empirical evidence. There are two major implications of the second-person approach to social neuroscience which can be derived from the description of the approach in the previous chapter. The first major implication concerns the 'least reducible unit' of social knowledge. While cognitivist accounts of social cognition argue that social cognition can be reduced to representations in the minds of two individuals, the second-person approach advocates the view that social interaction cannot be reduced to the content of single minds but that knowledge resides in the interaction 'between' two agents (Schilbach et al., *in press)*. Experimental data supporting this claim in its strictest interpretation would have to demonstrate that knowledge about the other can emerge from the interaction process *per se* and that the emergence of this knowledge can *exclusively* be explained by the dynamics of the interaction.

Indeed, the epistemic value of interaction dynamics has been demonstrated for the first time using an innovative 'perceptual crossing paradigm' (Auvray et al., 2009). The central question of this experiment was whether a human interaction partner can be recognized based on intrinsic properties of shared perceptual activity in a minimalist tactile environment. To this end, two blindfolded participants interacted by moving their mouse

cursor in a one-dimensional space on a computer screen. In this space, they encountered a fixed and a continuously moving object, an object representing the other's cursor, and a 'shadow' of this object. At each encounter participants received a tactile stimulation, and were instructed to click the mouse whenever they thought a particular stimulation is due to an encounter with the other. There was one crucial difference between the cursor and the shadow. Both objects moved exactly in the same way, but participants only received simultaneous tactile stimulation when their avatars encountered each other. Thus, only when they met, both were stimulated simultaneously. Results demonstrated a dissociation between task performance and awareness: A distinction between the avatar and the shadow was only detected in the absolute number of encounters, but not in the relative frequency of clicks. This suggests that participants were unable to recognize the other deliberately, but distinguished the avatar and the shadow unconsciously "as a consequence of the mutual search for one another which make the encounters between the participants far more frequent" (Auvray and Rohde, 2012, p. 2). Although the universality of these findings needs to be confirmed by research in other domains of social interaction, this simple paradigm has provided a first demonstration of meaning emerging from the interaction process itself. Importantly, however, the second-person approach does not stand or fall with interaction dynamics. Sufficient support for this approach would also be provided by demonstrations that the understanding of a person hinges on the observation of the effects our actions have on this person (Wolpert et al., 2003).

The second implication concerns the neural mechanisms supporting online social interaction. It is unlikely that entirely separate brain systems subserve offline and online social cognition, as both modes are directed at understanding other persons. However, if offline and online interaction *were* fundamentally different operational modes, differences in the underlying neural systems should be observable. For example, it is possible that discrete temporal and anatomical couplings of regions originally associated with the MNS and the SBN – and therefore with offline processes – mediate online social cognition. While there are numerous studies investigating offline social cognition either from a first- or third-person perspective (» Chapter 2.1), and a growing number of empirical approaches to study minds in interaction (Schilbach et al., *in press*), studies directly comparing observation and interaction are still lacking. Unfortunately, it is therefore impossible at present to reach any conclusion about *specific* neural mechanisms supporting online social interaction.

2.2.2.1 Empirical Approaches to the Neural Mechanisms of Online Social Cognition

This section introduces existing approaches to the study of online social cognition. Figure 2 is an adaptation of a graph by Schilbach et al. (*in press*) which demarcates the experimental landscape of neuroscientific research on social interaction. The grey scale of the elements of Figure 2 indicates the depth of empirical coverage of the respective categories with darker shades of grey indicating lesser coverage.

Category 1 comprises studies comparing direct social engagement and detachment by employing comparably static social stimuli. Among these are facial expressions, movements, or gestures which are either directed at the participant or at another person. Overall, this type of studies demonstrated across a range of stimulus categories that in comparison to detachment, direct engagement recruits the SBN (e.g. Bristow et al., 2007; Conty and Grèzes, 2012; Kuzmanovic et al., 2009; Tylén et al., 2012) and the amygdala (Schilbach et al., 2006). Particularly the finding of differential amygdala activation suggests that there is increased emotional engagement when facing self- versus other directed actions. *Category 2* is essentially an artificial category which has emerged as a by-product of the category system by Schilbach and colleagues (*in press*) and shall therefore not be discussed here.

Category 3 comprises paradigms in which participants are able to observe the effects of their behavior on another person. Although such structured interactions are principally one-way interactions and have been disparaged as 'pseudo-interactive' (Konvalinka and Roepstorff, 2012) they capture one of the most crucial aspects of social interaction: They enable participants to experience feedback to their actions and thereby to adjust their behavior based on the reactions they are facing. Interactions of this category usually involve an *initiator* and a *responder*, such as it is the case in interactive eye-tracking paradigms. They either allow participants to interact with others via live video-streams (Redcay et al., 2010), or with virtual agents in a gaze-contingent fashion (Wilms et al., 2010). The application of these paradigms will be discussed in the section on social gaze (» Chapter 3). Economic games – such as prisoner's dilemma, ultimatum, or trust games – represent another class of structured social interactions which are characterized by highly formalized turn-taking procedures (Rilling and Sanfey, 2011).

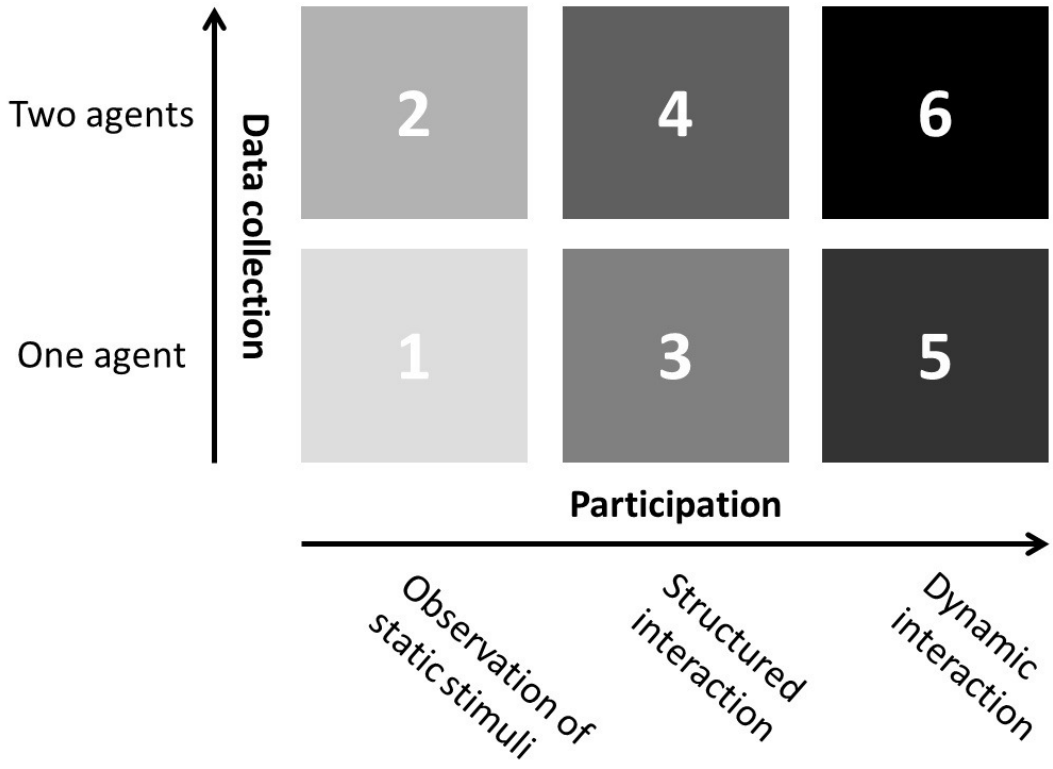

Figure 2. Landscape of empirical research on social interaction. This diagram depicts different categories of experiments addressing the cognitive processes and neural mechanisms of online social interaction (adapted from Schilbach et al, *in press*).

Category 4 involves the same paradigms as category 3. The difference is that data are collected from *both* interaction partners while they engage in a structured social interaction. This can be achieved using so-called hyperscanning paradigms, in which the brain activity of two (or more) persons is measured simultaneously while they are engaged in social interaction (Babiloni et al., 2006 (EEG); Baess et al., 2012 (MEG); Montague et al., 2002 (fMRI)). Despite the technical possibility of hyperscanning, the few existing studies so far have largely relied on game theoretic experiments (Fliessbach et al., 2007; King-Casas et al., 2005; Tomlin et al., 2006) and hence might fail to capture the dynamics of real social interaction as well as the ecological validity with respect to real-life social interactions (Konvalinka and Roepstorff, 2012). However, there are novel advances towards realistic, truly interactive eye-tracking paradigms involving two participants (» Chapter 3).

Categories 5 and 6 can effectively be grouped together. In exceptional cases – for example when the connection of two MRI scanners for hyperscanning is impossible due to technical or financial restraints – it might make sense to create a truly interactive situation in which data of only one of the participants are collected. In most cases, however, the application of an interactive setup for realistic dyadic interactions will demand collecting the data of both participants, in particular because the analyses of individual data sets are interdependent. As the black box depicting category 6 (Figure 2) illustrates, there are barely any studies investigating unconstrained social interactions. Until present, there is only one dual EEG study investigating two interactors spontaneously imitating their hand movements without *any* further instructions (Dumas et al., 2010). Interestingly, these authors found that behavioral synchrony is associated with a coupling of oscillatory brain activity in the alpha-mu rhythm. The source of this neuronal synchrony could be localized approximately to the TPJ. Furthermore, the alpha-mu rhythm has been related to activity in the MNS and has previously been described as an oscillatory marker of social coordination (Tognoli et al., 2007). Although it has recently been made possible to combine dual eye-tracking and hyperscanning (Saito et al., 2010), there is not yet a study employing such a setup in an ecologically valid manner. A main problem here is the lack of experimental control and the choice of the variables entering data analysis. Whilst not offering fully unconstrained interactions, structured interaction paradigms have an advantage in this respect.

In summary, a variety of approaches to the study of online social cognition have been proposed over the last years. While it is obvious that state-of-the-art approaches have surpassed categories 1 and 2, research on entirely unrestrained social interaction is still in its infancy and faces several financial, technical, and methodological problems. For instance, a major conceptual issue for social neuroscience concerns the question whether fMRI hyperscanning is altogether suited to capture the neural signature of interacting brains. One reason is the limited space inside the scanner bore and the susceptibility to movement artifacts. The main reason, however, is the low temporal resolution of the blood-oxygen level response (BOLD) underlying fMRI which might render a measurement of the neural processes supporting minute interactions of mutually coordinated agents impossible (Konvalinka and Roepstorff, 2012). Since the potentials of fully interactive paradigms cannot presently be exhausted, structured interaction paradigms falling into categories 3 and 4 appear most auspicious at the moment. This is especially true because there are still many

open questions concerning the neural mechanisms related to the subjective experience of being engaged in social interaction in real-time. For this reason, the studies comprising this thesis rely on structured interactions. The novelty is that these interactions are operationalized in the domain of social gaze, which endows them with greater ecological validity than previous paradigms.

2.2.3 The Rewarding Nature of Social Encounters

The previous sections have focused on the participatory aspects of social interaction and ways to address them. Another essential feature of online interactions is emotional engagement (Schilbach et al., *in press*). Through a wide range of disciplines, it has been suggested that social interaction is intrinsically rewarding (e.g. Baumeister and Leary, 1995; Krach et al., 2010; Moll and Tomasello, 2007). Recently, Lebreton and collaborators (2009) linked personality traits with neuroanatomical measurements and demonstrated that social reward dependence positively correlates with grey matter increases in the orbitofrontal cortex and the basal ganglia, including the ventral striatum. As social reward dependence provides a measure of an individual's propensity to engage in social interaction (Cloninger et al., 1993), these results suggest a direct link between the reward system and the motivation to participate in social interaction with others.

Indeed, activation of reward-related neurocircuitry was demonstrated in previous neuroimaging studies on social interaction. Rilling and colleagues scanned participants playing a Prisoner's Dilemma Game with another person or a computer (Rilling et al., 2002). In each round, players could choose to defect or to cooperate, whereby cooperation represented the riskier choice in terms of monetary outcome. Mutual cooperation led to increased activity in the mOFC, ACC, and the VS when interacting with a human. The VS was not activated during mutual cooperation with a computer, thus suggesting that cooperation with a human but not a computer correlates with positively reinforcing activity in this area. Also the intentions to trust someone (King-Casas et al., 2005) as well as someone's reputation for positive reciprocity during multi-round trust games are encoded in the VS (Phan et al., 2010). Another study compared neural activity when participants played a competitive video game against human or computer opponents (Kätsyri et al., 2012). Results indicated that winning led to greater activity in the striatum than losing. This activity

was further enhanced when participants thought they were interacting with a human opponent, again suggesting a specific role for reward-related processes during social interaction. Although these studies assess the neural mechanisms supporting online social interaction, there are two reasons why they do not answer the question concerning the reward value of social interaction *per se*. Firstly, most of them have relied on game-theoretical approaches (Rilling and Sanfey, 2011). These approaches involve comparatively sophisticated interactions requiring the build-up of trust or reputation in turn-taking games and hence comprise high-level social concepts such as cooperation, trust, fairness, or altruism. Secondly, in other studies applying a human-computer distinction subjects were informed in advance whether they were going to interact with another person or a computer program (e.g. Decety et al., 2004; Gallagher et al., 2002; Kätsyri et al., 2012; Kircher et al., 2009; Rilling et al., 2004; Sanfey et al., 2003). One advantage of the non-verbal Turing test presented in ***studies 3 and 4*** is that the interactions are of a basic nature but nonetheless ecologically valid. Furthermore, participants are not pointed towards the nature of their interaction partner in advance, but develop a subjective experience of being in social interaction with another human through the interaction process itself.

3 Social Gaze: A Non-Verbal Cue System Suited for Studying Online Interaction

In the present thesis, social gaze behavior has been selected as a means to operationalize online interactions for two major reasons: Firstly, gaze provides a salient non-verbal cue system commonly used in everyday social encounters and combines perception and action in single acts of looking. Secondly, gaze can be measured in combination with neural activity by using fMRI-compatible eye-tracking devices. The application of MRI has the advantage of providing access to medial and orbital cortical as well as subcortical regions, such as the amygdala and the basal ganglia. These are critically involved in affective and motivational aspects of social interaction and not easily accessible using EEG or MEG.

Gaze serves a variety of social-cognitive functions beyond mere visual detection. It has been related to information seeking, signaling interpersonal attitudes, regulating the synchronicity of speech during dialogues, and plays a significant role in the regulation of interpersonal distance as well as the avoidance of undue intimacy (Argyle et al., 1973). Furthermore, the eye region provides social information related to a person's identity, emotional state and focus of attention. Consequently, whenever we look at a face, the eyes are the primary and most consistent target of our visual attention (Haith et al., 1977; Walker-Smith et al., 1977). Evolutionary biologists found the human eye to have a unique morphology (Kobayashi and Kohshima, 2001). Non-human primates generally have darkened eyes, which makes it difficult to distinguish the iris from the surrounding sclera. In contrast, the human eye is characterized by a largely exposed, depigmentated sclera surrounding a dark-colored iris. The development of these features facilitates the detection of the gaze direction of other individuals and has been paralleled by the development of brain mechanisms supporting social cognition, thereby suggesting that gaze is a crucial component of our social-cognitive skills (Emery, 2000).

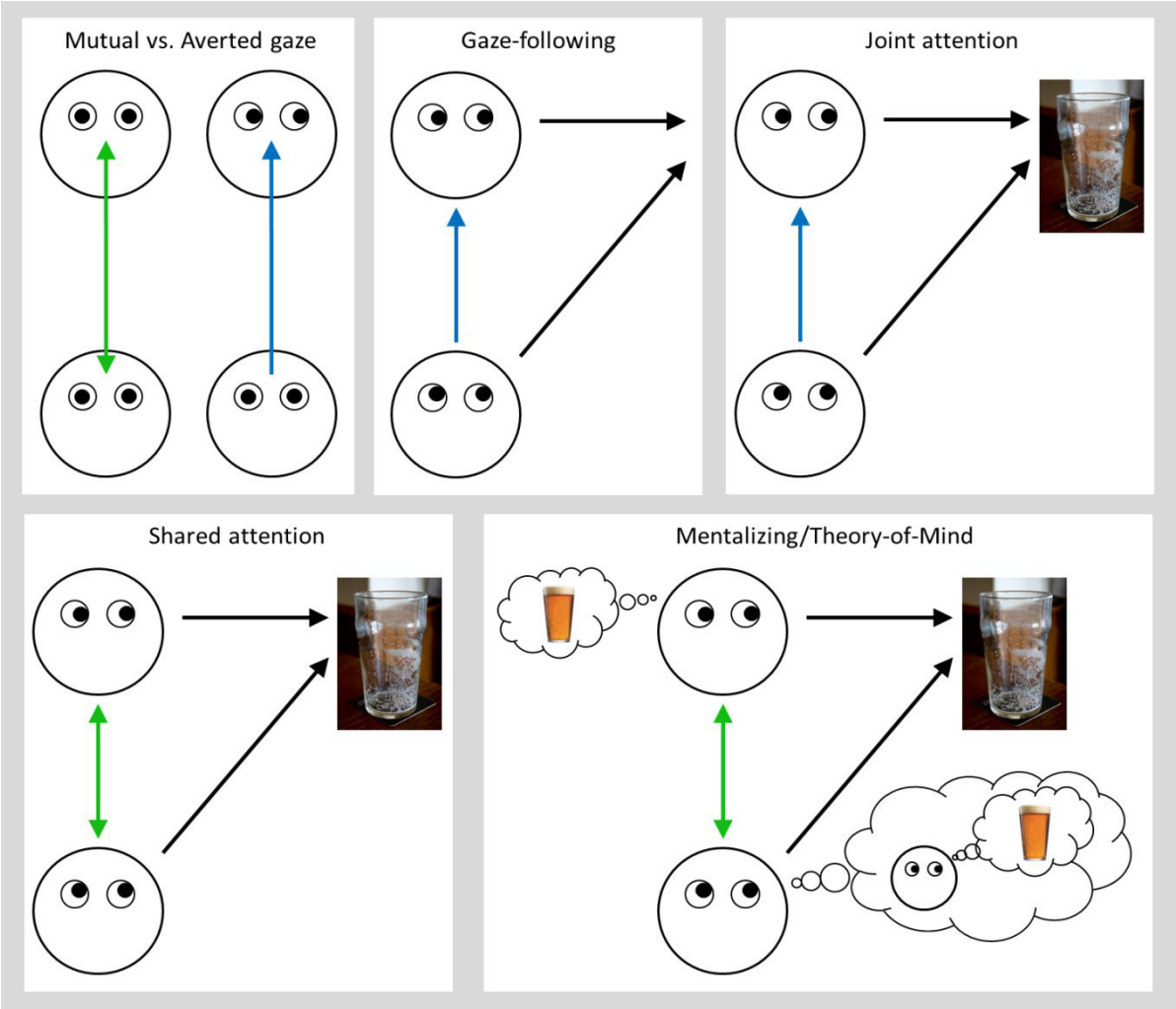

Figure 3. Core processes of social gaze. Green arrows indicate mutual awareness, blue arrows indicate unidirectional awareness (adapted from Emery, 2000).

3.1 The Core Processes of Social Gaze Behavior

Emery (2000) has identified direct (or mutual) gaze, gaze aversion, gaze-following, joint attention, and shared attention as core processes constituting social gaze behavior (Figure 3). The discussion of these processes is necessary to comprehend how gaze behavior can be used to construct meaningful social interactions. Notably, gaze has an affective as well as a communicative dimension. The following sections will only briefly touch upon the affective aspects of gaze (cf. Itier and Batty, 2009) as this thesis focuses on gaze as a cue system supporting non-verbal social interaction. The core processes and their neural underpinnings will be discussed here.

3.1.1 Opening the Communicative Channel by Mutual Gaze

Any gaze-based interaction starts with two individuals looking at each other, a situation referred to as *mutual gaze (*or *direct gaze*). In contrast to other species, direct gaze is not (necessarily) perceived as a threat in humans but serves as a natural attractor of attention already in newborns (Farroni et al., 2002). Numerous studies have found that direct gaze is a powerful modulator of cognition – a phenomenon referred to as the 'eye-contact effect' (Senju and Johnson, 2009). For example, it has been shown that faces displaying direct gaze are spotted faster among distractor items and memorized better than faces with averted gaze. Specific facial features and facial identity are detected earlier in faces expressing direct gaze. Direct gaze also exerts a positive influence on our judgment of other's attractiveness (Mason et al., 2005) and increases the likeability of objects that are presented adjacent to faces expressing direct gaze (Strick et al., 2008). Most importantly, however, it has been described as an ostensive cue which signals that a behaviorally relevant act is going to follow (Csibra and Gergely, 2009). Direct gaze thus serves a particular behavioral function: It indicates communicative intent and thereby 'opens the channel' for any meaningful social interaction (Cary, 1978).

A recent model by Senju and Johnson (2009) proposes that the eye contact effect is brought about by a subcortical 'fast-lane' via the amygdala and low-level visual areas including the superior colliculus and the pulvinar. Such a subcortical shortcut involving the amygdala has, for instance, also been demonstrated for the processing of fearful stimuli (LeDoux, 2000), which suggests that subcortical mechanisms generally facilitate the processing of behaviorally significant stimuli. In the case of direct gaze, this mechanism is supposed to modulate social brain areas involved in the detection of gaze direction (aSTS) and intentionality (mPFC, pSTS) depending on the present context and task demands (e.g. visual search, detection of facial identity etc.) which are administrated by the dlPFC.

3.1.2 The Perception of Averted Gaze and Gaze-Following Behavior

The perception of *gaze aversion* is a prominent cue signaling that another individual's attention is not directed at us but at another aspect of the environment. Many studies using Posner-task-like gaze cueing paradigms have demonstrated that the detection of *averted gaze* results in a shift of visuo-spatial attention by means of reflexive *gaze-following*

behavior even when a gaze cue is counter-predictive of a target (cf. Shepherd, 2010). The development of gaze-following is also considered a critical hallmark of social-cognitive development, as it relates to an understanding of the field of view of another individual and the relevance this individual assigns to objects in that field. There are also affective differences between direct and averted gaze. While facing direct gaze has positive effects on likeability and attractiveness ratings, experiencing averted gaze can result in feelings of social exclusion and reduced self-esteem (Wirth et al., 2010). One study measured participants' brain activity while they watched video sequences in which an animated character walked toward them and either expressed mutual or averted gaze when passing them (Pelphrey et al., 2004b). Mutual gaze led to greater activation of the pSTS than averted gaze, which suggests that this region encodes approach and avoidance associated with direct and averted gaze, which might play a role in the detection of communicative intent. Interestingly, an electroencephalographic (EEG) study showed that observing faces displaying direct and averted gaze was associated with patterns of EEG activation related to approach and avoidance systems, respectively (Hietanen et al., 2008). This has led to the general idea that "more gaze tends to elicit more positivity" (Wirth et al., 2010, p. 869) in object and person perception.

The detection of gaze direction is supported by activation of the STS, which is commonly separated into an anterior and a posterior region. Neuroimaging studies in humans have consistently reported the pSTS to encode the perceived gaze direction of other individuals (cf. Nummenmaa and Calder, 2009). However, in recent years it was shown that also the aSTS encodes particular gaze directions in a more fine-grained fashion than the pSTS (Carlin and Calder, *in press*). While the pSTS responds rather generally to the direction of attention and is not only activated by perceived gaze direction but also by head direction and body posture (Redcay, 2008), the aSTS contains neural populations encoding gaze direction independently of head direction and physical features of faces (Carlin et al., 2011, 2012). Carlin and Calder (*in press*) suggest that there is a posterior-to-anterior specialization for gaze direction which is increasingly invariant to gaze-irrelevant features. Similar hierarchies with increasingly anterior specializations have been found for the coding of facial identity in cells of the macaque temporal cortex, suggesting that "hierarchical progressions toward view invariance may therefore be a general property of high-level face representations" (Carlin et al., 2011 p. 1820). Evidence that the pSTS is sensitive to the

social context in which a gaze shift occurs comes from an experiment in which participants watched brief video clips in which a visual stimulus appeared to the left or right side of a virtual agent looking at them. The agent then either shifted its gaze towards the stimulus or towards the blank side. Whenever the agent's response was incongruent, activity in the right pSTS was increased. As a visual stimulus should evoke more interest than a blank space, the shift away from the stimulus violated participants' expectations regarding the agent's behavior. It was therefore concluded that the pSTS might be involved in detecting intentions expressed by gaze shifts (Pelphrey et al., 2003). In a similar study, the manipulation did not involve gaze shifts, but observed reach-to-grasp movements towards or away from a visual target (Pelphrey et al., 2004a). Activity in the pSTS was also increased for incongruent movements, thereby implicating a general role in analyzing the intentionality of actions.

Given the contribution of the pSTS to multiple different aspects of gaze processing (approach-avoidance, general detection of social attention, encoding of intentions…) it is difficult to postulate a common function of this region. This is further aggravated by the involvement of the STS in speech perception, audio-visual integration, and the analysis of biological motion. It has thus been proposed that the function of the STS critically depends on network co-activations and that future studies should focus on connectivity analyses to detect commonalities and differences in STS activity between different tasks (Hein and Knight, 2008).

3.1.3 Joining and Sharing Attention

Following another individual's gaze to a novel focus of visual attention creates a situation of *joint attention (JA)*. In contrast to a situation in which two individuals coincidentally look at the same object, the follower is aware of sharing the gazer's focus of attention in joint attention. According to the definition of Emery (2000), however, JA does not require the gazer to be aware of the gaze-follower's reaction. A distinction has therefore been made between *joint* and *shared attention (SA)*, with the latter requiring *both* individuals to be aware of focusing on the same object *and* on each other. Whereas the terms JA and SA are often used interchangeably in the literature (and usually subsumed under the label JA), the comparably modular and clear-cut definitions of core processes of social gaze of Emery (2000) shall be adopted here. They serve as the conceptual basis for an empirical

investigation of potential behavioral differences between JA and SA in ***study 2***. For sake of simplicity, however, the term JA will be used in the following.

While mutual gaze and gaze-following represent *dyadic* processes involving two individuals, JA represents a *triadic* interaction involving a 'referential triangle' of two individuals and some third entity (e.g. object, person, location etc.) in the environment (Carpenter et al., 1998). Considering that people look where they attend and where they intend to act, JA is hence considered fundamental to an understanding of other minds. An important distinction is made between responding to other people's offerings of JA and initiating JA (Mundy and Newell, 2007). While responding to a bid for joint attention by following someone's gaze can also be observed in non-human primates, the initiation of JA is believed to represent a uniquely human capacity. Indeed, it has been demonstrated that starting at the age of 9 months, children show an intrinsic motivation to actively share attention with parents and caretakers by directing their gaze to an object, and to reengage them in social interactions which have previously been interrupted (see Tomasello et al., 2005 and references therein). JA thereby combines an intrinsic motivation to share with the establishment of perceptual common ground in a reciprocal fashion (Clark, 1996). This has been argued to create "a shared space of common psychological ground that enables everything from collaborative activities with shared goals to human-style cooperative communication" (Tomasello and Carpenter, 2007, p. 121). In sum, JA is the origin of any meaningful non-verbal social interaction and therein provides a tool to study engagement and participation in online social interaction.

Neuroimaging literature on JA is still comparatively sparse. Due to the lack of interactive technologies, researchers were facing the problem that they could only address participant's responses to bids of JA by using *gaze-cueing* paradigms. Materna et al. (2008) found that engaging in JA by following another individual's gaze cue to an object in space recruited the pSTS, which is consistent with studies on the perception of gaze direction. More importantly, these authors demonstrated that this activation is specific for gaze cues, while symbolic cues recruit the IPS, which is more generally implicated in re-directing spatial attention. Another study reported activity in the mPFC while participants engaged in JA (Williams et al., 2005), thereby suggesting a link between JA and mentalizing. Although these results are interesting, the methods used did neither capture the reciprocal nature of JA,

nor did they allow for disentangling differences in the neural mechanisms subserving initiating and responding to JA.

The development of the interactive setup in ***study 1*** made it possible to use interactive eye-tracking in combination with MRI for a systematic study of responding to and initiating JA and NJA (Schilbach et al., 2010). Their results indicated that the mPFC is activated preferentially during the response to JA, whereas the successful initiation of JA specifically recruits the ventral striatum (VS) which is a central component of the reward system. This study provided first-time evidence for an intrinsic motivational mechanism to share attention and thereby, in a broader sense, initiate social interaction with others (Tomasello et al., 2005). In another study using live video feeds, Redcay and colleagues compared self- and other-initiated JA using a paradigm in which participants were instructed to play a game with the experimenter (Redcay et al., 2012). They had to locate an object on the screen following the gaze cue of the experimenter, or to help the experimenter to find the object by providing them with a gaze cue. A solo attention condition served as a control condition. The most reliable differential activation for JA was observed in the right pSTS. This is consistent with previous studies that have demonstrated the involvement of this region in other-initiated JA (Materna et al., 2008; Redcay et al., 2012). As it has been described above, the pSTS is crucially involved in decoding the direction of another person's gaze direction as well as the intention behind a gaze shift (Nummenmaa and Calder, 2009). Both self- and other-initiated JA showed greater activation of the dorsal mPFC, which is consistent with the results of previous studies (Schilbach et al., 2010; Williams et al., 2005) and is likely to reflect a sharing of attention that is necessary for an inference of another individual's mental states. During other-initiated JA, however, the ventral mPFC was activated to a greater extent as in self-initiated JA. This activation extended in the mOFC, which is considered part of the brain's reward system. As the gaze cue of the experimenter was a cooperative cue in the framework of a collaborative game involving a joint intention (i.e. 'finding the object together'), the authors argued that this activation might be related to the anticipation of rewards. However, a recent meta-analysis demonstrated that it is specifically involved in the subjective hedonic experience of reward (Peters and Büchel, 2010). This suggests that this activation is related to the experience of receiving a cooperative cue from another person rather than to reward anticipation. Self-initiated JA recruited a fronto-parietal attention network (FPAN), which is explained by

greater attentional demands attributed to the voluntary initiation of JA in comparison to following someone's gaze. In the study by Schilbach and colleagues (2010), this network was only recruited during NJA. Although this discrepancy cannot be clarified here entirely, it might be related to the explicit instruction to engage in NJA that participants received in the study by Schilbach and co-workers (2010), while the control condition in the study using live video feeds was a solo attention condition in which the experimenter disengaged from the interaction by closing her eyes.

In sum, JA is associated with activation of the mPFC, pSTS, and networks more generally related to attentional processing. Especially the mPFC has been activated reliably during JA in most studies and thereby appears to represent a neural mechanism at the intersection between gaze behavior and an understanding of intentions in triadic relations. Considering the prominent role of the mPFC in offline mentalizing studies (» 2.1.2.1), these findings suggest that the mPFC broadly supports the understanding of others by enabling a 'meeting of minds' (Amodio and Frith, 2006). In contrast, the reward-related activation of the VS during self-initiated JA has only been found in a single study so far (Schilbach et al., 2010). Consequently, the exact nature of reward mechanisms during gaze-based interactions was a central issue of investigation in ***study 4*** of this thesis.

4 Interacting with Virtual Agents

Virtual reality techniques have found their way into numerous areas of neuroscientific research (Bohil et al., 2011). In social psychology and social-cognitive neuroscience, the use of virtual agents is promoted as they allow isolating a cue of interest, such as gaze reactions, while neutralizing confounding variables (Fox et al., 2009). Especially in neuroimaging studies, a reduction of the 'band width' of social interaction is essential in order to maintain experimental control. Recent research on so-called *anthropomorphic* virtual agents – which have realistic human features while still being easily recognized as artificial – has suggested that they provide an excellent tool to study online social interaction (cf. Vogeley and Bente, 2010).

In all studies comprising this thesis, participants engage in reciprocal interactions in which they are directly addressed by a virtual agent who reacts to their gaze behavior in a gaze-contingent fashion. At the outset of this thesis, this approach was an absolute novelty. In the meantime, two other groups have developed interactive eye-tracking setups to study real-time interaction. One setup involves the use of virtual characters similar to the one presented in ***study 1*** (Grynszpan et al., 2011, 2012), while the other allows face-to-face interaction between a subject inside an MRI scanner and an experimenter via a live video-feed (Redcay et al., 2010). Both approaches have advantages and disadvantages. Obviously, the use of video-feeds allows the most realistic kind of gaze-based interaction. It has been shown that watching live faces but not pictures of faces increased left-frontal EEG asymmetries and arousal related to approach behavior (Pönkänen et al., 2008) which has been attributed to an enhancement of mentalizing and self-awareness induced by the interaction with a real face (Pönkänen et al., 2011). Accordingly, observational studies found that when participants watched movie clips of real humans versus animated humans performing actions, there was enhanced activity in areas of the SBN, such as the mPFC, TPJ, and STS (Han et al., 2005; Mar et al., 2007). Notably, it is possible that the advantage that has been attributed to using live faces might be partially due to the nature of the facial stimuli used in the studies cited above. For example, Han and colleagues (2005) used excerpts of TV cartoons which lack a variety of facial features whose presence might be

important for social-cognitive processes. Conversely, in another study, real scenes were interleaved with cartoon-rendered scenes of the same movie that were still extremely realistic (Mar et al., 2007), hence possibly triggering the 'uncanny valley effect', an unsettling experience that emerges when some entity appears extremely human without actually being human (Mori, 1970).

The investigation of gaze behavior in interactions with live faces is prone to substantial noise because gaze direction is only one – albeit powerful – out of multiple facial cues conveying non-verbal information (Grant, 1969). Even slight changes in the muscles controlling the eye region or the labial angles can result in differences in the perception of a person (Ekman and Oster, 1979) and it has often been demonstrated that facial features and facial configuration have a general bias on our impression of others (Berry and McArthur, 1986). Studies focusing on gaze-based social interaction hence require strict control of these features which is hardly possible when an experimenter or actor functions as the interaction partner. These problems can be circumvented by using carefully designed anthropomorphic virtual agents. The interaction with such agents results in comparable reactions and social behaviors as the interaction with real humans. For example, socially induced inhibition as well as facilitation of task performance could not only be observed in the presence of another person, but also in the presence of human-controlled avatars (Hoyt et al., 2003). Likewise, participants' regulation of interpersonal distance and approach behavior is comparable in immersive virtual environments and real social encounters (Bailenson et al., 2003). It was also demonstrated that participants displayed empathic concern for virtual agents in distressful situations, thereby demonstrating that encounters with virtual characters readily elicit prosocial behavior (Gillath et al., 2008). Finally, a study by Bente and collaborators revealed that the simulation of gaze behavior by virtual characters in face-to-face interactions results in similar experiences of social presence and intimateness as real gaze behavior (Bente et al., 2007).

To conclude, the use of live video feeds principally allows a holistic examination of the neural correlates of face-to-face interactions. However, with presently available neuroimaging methods it is difficult to correlate neural activity with specific components of complex social interactions. It is therefore beneficial to adopt a 'divide-and-conquer' strategy and decompose these interactions into meaningful constituents. Virtual agents

permit investigating the effects of gaze reactions while other facial features such as emotional expressions or attractiveness are kept constant. These features can then be modulated in a step-by-step fashion, for instance by pairing congruent gaze reactions with a smile, a frown, or an eye-blink. Taken together, the decision between real and virtually mediated interactions is presently a decision between a bottom-up and a top-down approach to the study of social interaction.

5 Research Agenda of this Thesis

The aim of this thesis was to create an experimental paradigm which would allow to investigate the subjective experience and motivational basis of being engaged in online social interaction – i.e. face-to-face and in real-time – both on a behavioral and neural level. The construction of a realistic gaze-based interaction paradigm compatible with fMRI required several steps. The first step consisted of the development of an interactive eye-tracking paradigm in which participants can initiate interactions with anthropomorphic virtual characters in a gaze-contingent fashion. The challenge of ***study 1*** was to design a paradigm in such a way that it could be used with on any hardware with a variety of eye-trackers to allow for behavioral testing as well as application in an MRI scanner. Before a gaze based interaction paradigm can be crafted, it is necessary to explore the dynamic and temporal aspects of social gaze behavior on the basis of single gaze trials. ***Study 2*** aimed at uncovering the latencies of another individual's gaze-following reactions that were experienced as contingent upon participant's own gaze shifts. These 'natural' latencies were then used in another experiment addressing differences in the dynamics of gaze shifts in the conceptually distinct processes of JA and SA. It was hypothesized that the initiation of JA would require less gaze shifts between the virtual agent and an object on the screen than the initiation of SA. Using the information obtained in ***study 2***, the actual interaction paradigm, which represents the core of this thesis was developed in ***study 3***. The objective was to design a task that would allow differentiating participants' subjective experience of interactions as social or as non-social. To this end, a non-verbal Turing test was generated in which participants had to judge whether a virtual agent had been controlled by a human conspecific or a computer algorithm based on the gaze reactions of this agent during brief interaction sequences. The interaction partner was introduced either as naïve to the participants' task, as explicitly cooperative, or as openly competitive. This allowed assessing the patterns of gaze reactions participants judged as indicative of real human interaction in different types of interaction contexts. Finally, ***study 4*** was a within-subject version of the non-verbal Turing test adapted for fMRI which targeted the intrinsic motivational foundation and rewarding nature of social interaction.

5.1 Own Experimental Studies

5.1.1 Study 1: It's in your eyes – using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. (Wilms M, Schilbach L, **Pfeiffer UJ**, Bente G, Fink GR, Vogeley K, 2010. *Social Cognitive and Affective Neuroscience*, 5(1), pp. 98 – 107.)

In this study we sought to develop an interactive eye-tracking setup, which allows participants to interact with an anthropomorphic virtual agent on a computer screen via their eye-movements. An interactive paradigm means that that users cannot only react to the agent, but are also able to observe the agent moving its eyes in a gaze-contingent fashion, that is, in response to their own gaze behavior. Although gaze-contingent eye-tracking had been used in human-computer interaction research before the beginning of this thesis (Duchowski et al., 2004), no methods allowing reciprocal, non-verbal interaction in the context of social neuroscience existed at that time. The challenge of our approach was to create an MRI-compatible eye-tracking setup which extracts the gaze input of a user to control visual stimulation in real-time according to preset task definitions.

The algorithm controlling the interactive paradigm was originally implemented in Presentation™ (Neurobehavioral Systems, www.neurobs.com), a software package dedicated to visual stimulus presentation in psychological experiments. However, the algorithm is universally applicable and can therefore be used across a variety of programming languages and eye-tracking systems. Figure 4 (adapted from Pfeiffer and Weidner, 2013; Wilms et al., 2010) depicts the algorithm in a flowchart. Input and output are depicted by green flowchart elements. The algorithm enables both *overt* and *covert* feedback, the former being symbolized by blue and the latter by red elements. The blue parts of the flowchart depict components relating to the detection of gaze coordinates and overt feedback. Raw gaze coordinates are fed into a continuously moving sliding window with a preset number of gaze positions. A *moving average* is calculated from this sliding window to ensure smooth movements of a gaze cursor in cases of overt feedback. Standard deviations of gaze coordinates are monitored in order to detect coherent gaze periods. After a preset number of moving averages, a gaze period is accepted as a fixation. When covert feedback is used, it is validated whether a fixation is within a region of interest (ROI). This is depicted

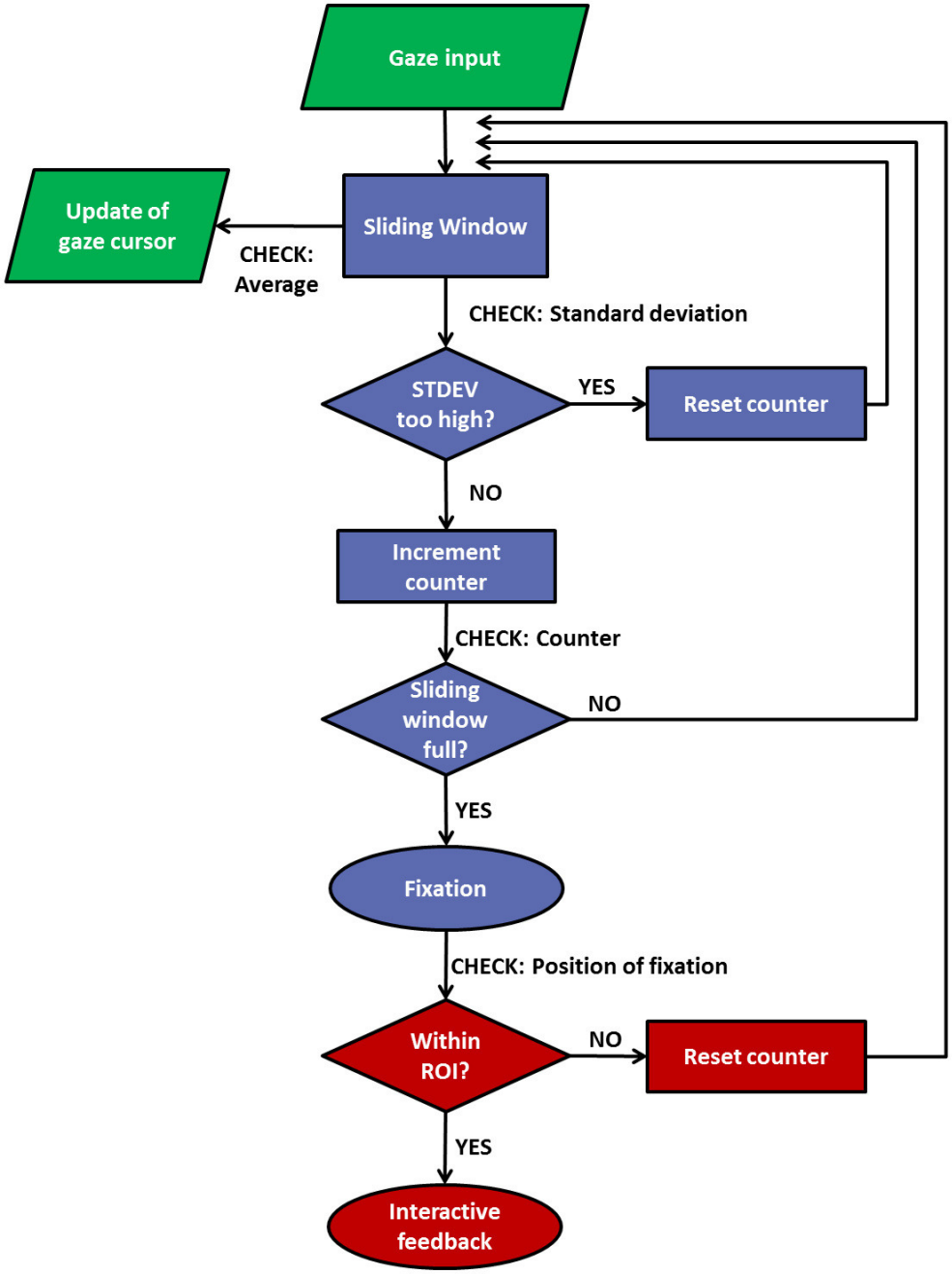

Figure 4. Interactive eye-tracking algorithm. Flowchart depicting the elements of the interactive eye-tracking paradigm established in study 1. Green elements denote in- and output, blue elements are important for both overt and covert feedback, and red elements relate to covert feedback (adapted from Pfeiffer and Weidner, 2013; Wilms et al, 2010).

by the red elements of the flowchart. Depending on the ROI fixated by a user, the visual stimulation on a screen can then be changed dynamically, for instance by making a virtual agent follow the user's gaze to engage in JA. Since the delay caused by the algorithm is in the range of few milliseconds (ms), gaze-based interactions can be simulated in real-time.

Taken together, interactive eye-tracking constitutes a significant enrichment of the toolbox of social neuroscience by allowing an assessment of *online* social cognition and enabling participants to observe an agent reacting to their own gaze shifts. This closes the 'perception-action loop' between two interacting individuals and allows investigations of initiating joint and shared attention as compared to merely responding to another person's gaze cues. The use of virtual agents provides the possibility to control for factors such as attractiveness and likeability of an interaction partner and permits focusing on gaze behavior in an ecologically valid fashion (e.g. Bailenson et al., 2005; Bente et al., 2007; Vogeley and Bente, 2010).

5.1.2 <u>Study 2:</u> Eyes on the mind: investigating the influence of gaze dynamics on the perception of others in real-rime social interaction. (**Pfeiffer UJ**, Schilbach L, Jording M, Timmermans B, Bente G, Vogeley K, 2012. *Frontiers in Psychology*, 3, p. 537.)

While previous studies have concentrated on responses to other persons' gaze behavior, the development of an interactive eye-tracking paradigm (***study 1***) made it possible to examine different aspects of *self-initiated* gaze behavior. Similar to other domains of non-verbal behavior (e.g. facial expressions, gestures…), gaze behavior has certain specific temporal and dynamic features. For example, when someone follows our gaze after a delay of one second it is likely to be a reaction to our gaze shift. In contrast, when there is a delay of ten seconds, we hardly experience the other's gaze as contingent upon our own. More precisely, two questions were addressed in this study: (1) How does the perception of JA depend on the congruency (i.e. following vs. non-following) and latency (i.e. temporal delay) of another individual's gaze reaction? (2) Is there a difference between gaze dynamics in JA and SA?

In each experiment of this study, participants interacted with a virtual agent in an adaption of the interactive eye-tracking setup designed in ***study 1*** and believed that the eye

movements of a confederate were visualized by the agent. In *experiment 1a*, they were asked to fixate one of two objects on the screen. The agent would then either follow their gaze or look to the other object. The latency of the agent's gaze reaction was varied from 0 to 4000 ms in steps of 400 ms. Subsequently, they were asked to rate how contingent they experienced this reaction on their gaze shift on a 4-item scale. Overall, participants rated congruent reactions as more contingent on their own gaze with ratings decreasing for latencies greater than 800 ms. Furthermore, when reactions had a latency of 0 ms, they were experienced as coincidence and hence rated as rather non-contingent. *Experiment 1b* repeated experiment 1a without the incongruent condition to increase participants' sensitivity to the timing of gaze reactions by decreasing the putative other's options to act. Starting at a latency of 400 ms, there was a highly significant linear decrease in relatedness ratings that was much more pronounced than in experiment 1a. Furthermore, an analysis directly comparing the congruent trials of experiments 1a and 1b demonstrated that the experience of contingency was earlier in experiment 1b than in experiment 1a (i.e. 400 vs. 800 ms). This suggests that the action possibilities of an interaction partner are implicitly taken into account when evaluating the contingency between our own and another person's gaze behavior. The findings of this study demonstrate that, depending on an interaction partner's options to act, gaze reactions with latencies between 400 and 800 ms are perceived as most contingent (i.e. most natural) upon our own gaze shifts. Moreover, given that latency hardly has an effect on contingency judgments when the other has another option to act (experiment 1a), our sense of agency is stronger in this situation although the causal link between our gaze shift is weaker than in experiment 1b.

In *experiment 2*, these 'natural' latencies between 400 and 800 ms were used to construct a paradigm in which participants' were either instructed to engage in JA or SA. In the JA group, they were instructed to respond as soon as *they themselves were aware that both they and the other directed their attention to the same object*. In the SA condition, they received the instruction to respond once *they were convinced that both of them were aware of each other directing their attention to the same object*. SA required more oscillatory gaze shifts between the agent and the object than JA. There was also substantially more inter-individual variance in the number of gaze shifts in SA. These findings show for the first time that JA and SA differ regarding the underlying gaze dynamics and can thus indeed be regarded as different processes (Emery, 2000). The great variance in the number of gaze

shifts in the SA group furthermore suggests that gaze dynamics might be an embodied correlate of mentalizing processes involved, which have previously also been reported to be subject to great inter-individual variance.

In sum, ***study 2*** sheds light on important dynamic features of self-initiated social gaze behavior that could not be observed previously due to a lack of appropriate method. The findings also aid a clear terminological distinction of core processes of social gaze. In particular, the hitherto theoretical distinction between JA and SA has been empirically supported. Finally, the results are of great significance for a realistic design of prospective studies on gaze-based interactions.

5.1.3 Study 3: A non-verbal Turing test: Differentiating mind from machine in gaze-based social interaction. (**Pfeiffer UJ**, Timmermans B, Bente G, Vogeley K, Schilbach L, 2011. *PLoS One*, 6(11), e27591.)

Social interactions come in various forms and often present themselves as very complex. To study *online* social interaction under standardized experimental conditions, it is therefore indispensable to develop paradigms in which communicative bandwidth is limited while ecological validity is high. Furthermore, prior to addressing high-level interactions, it is necessary to reveal how the subjective experience of being engaged in social interaction in real-time shapes the behavior of an individual. This requires a paradigm in which participants subjectively experience an interaction as social or non-social depending on the dynamics of the interaction. Several decades ago, Alan Turing has developed such a paradigm, the so-called 'Turing test' (Turing, 1950). In this test, a participant engages in written communication with another human being or a computer program via a computer screen. The task of the participant is to deduce the nature of his present interaction partner solely from the conversation. Although Turing was interested in the question whether and which machines can be ascribed human-like intelligence, his famous test provides an excellent scaffold for separating situations of social from non-social interaction without informing participants *a priori* about the nature of their interaction partner. We thus created an experimental procedure termed the 'non-verbal Turing test', in which participants had to judge whether they were interacting with a human conspecific or a computer in brief, gaze-based interactions.

Each interaction consisted of a series of six gaze trials in each of which participants had to establish mutual gaze with a virtual agent before shifting their gaze to one of two objects on the screen. Consequently, the agent either followed their gaze or averted its gaze to the other object on the screen, thereby engaging in JA or NJA, respectively. Participants were instructed that in each interaction block the agent could be controlled by another human participant or a computer. Unbeknownst to them, however, the alleged interaction partner was a confederate and the agent was always controlled by a computer algorithm to allow for a systematic variation of JA and NJA ranging from zero to six out of six possible instances of JA. After each block, they had to engage in a 'humanness ascription task' by indicating whether the agent had been controlled by the other participant or the computer, or, in other words, whether they had experienced this interaction as social or non-social. In a series of three experiments using a between-subject design, the confederate was introduced

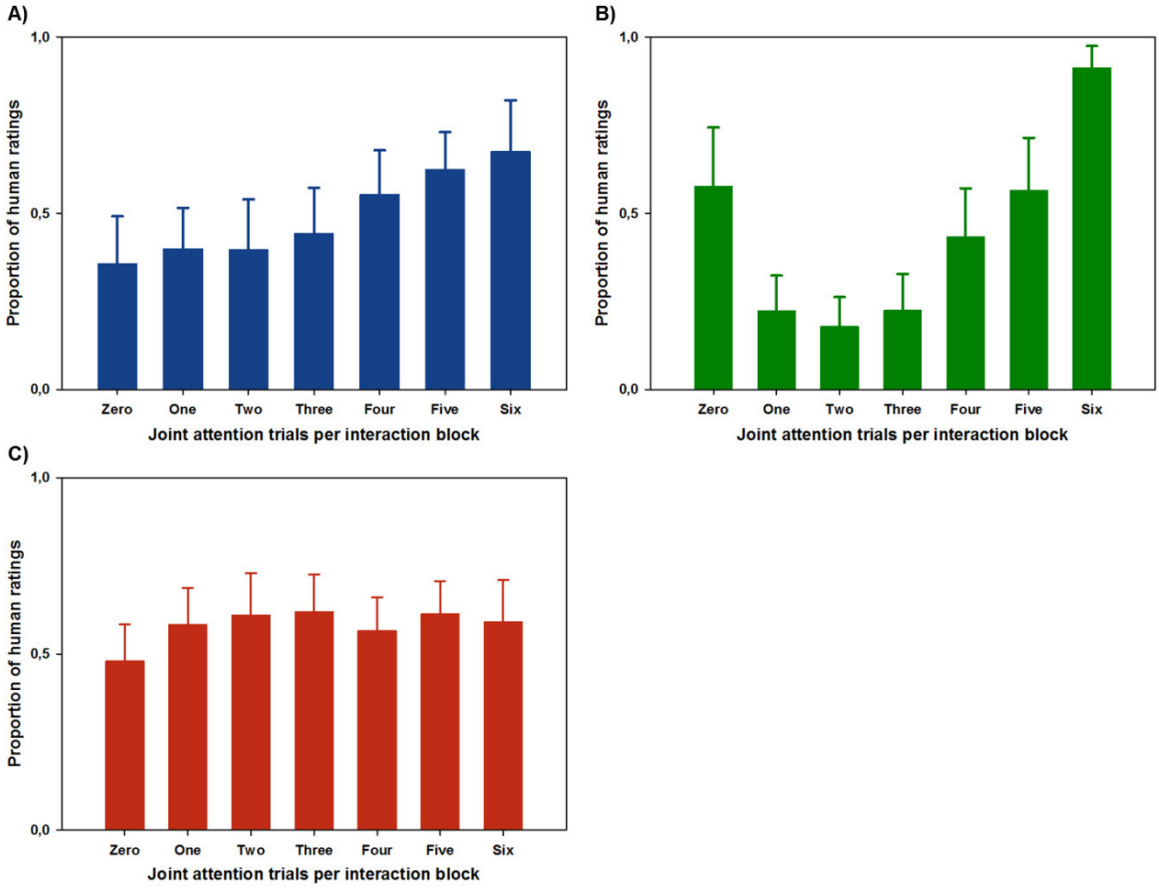

Figure 5. Humanness ascription in the non-verbal Turing test. These graphs depict the proportion of human ratings when the interaction partner was introduced as A) naïve, B) cooperative, and C) competitive (adapted from Pfeiffer et al, 2011).

as *naïve* to the participant's task, explicitly *cooperative*, or explicitly *competitive*. Results indicate that the ascription of humanness increases with higher degrees of gaze-following when participants interact with a putatively naïve partner (Figure 5A). In contrast, humanness was also ascribed in cases of high degrees of gaze aversion when the confederate had been introduced as cooperative, thereby indicating an expectation of contingent rather than merely congruent reactions (Figure 5B). In competitive interaction, neither congruency nor contingency of the agent's reactions played a role in humanness ascription (Figure 5C). In addition, the naïve and cooperative experiments were repeated using a concurrent cognitive load task to reveal the degree of automaticity of the different patterns of humanness ascription. The cognitive load tasks revealed that high-levels of JA are automatically considered as indicative of a human interaction partner, whereas the ascription of humanness in the cooperative condition is lacking such an automatic component, thereby indicating a comparably strategic integration of the other's reactions.

To begin with, participants readily believed that they were engaging in real social interactions implemented in a virtual environment. These results hence demonstrate that gaze behavior displayed by virtual agents can be applied to construct ecologically valid social interactions. Furthermore, the findings on humanness ascription show that humans have a default expectation of reciprocity in interaction. An interactor's presumed disposition to cooperate appears to drastically alter the evaluation of this interactor's reactions. Taken together, the non-verbal Turing test offers an appropriate tool to study the neural mechanisms associated with the experience of engagement in interaction (***study 4***).

5.1.4 <u>Study 4:</u> Why we interact: On the functional role of the ventral striatum during real-time social interactions. (**Pfeiffer UJ**, Schilbach L, Timmermans B, Kuzmanovic B, Georgescu AL, Bente G, Vogeley K., *Submitted*)

An intrinsic motivation for social interaction has often been proposed and is thought to be unique to the human species (Tomasello, 2009). Indeed, various neuroeconomic studies have found reward-related activity during social interactions. However, the application of economic games in the study of social interaction usually entails the investigation of specific high-level social concepts such as trust, fairness, or cooperation. In contrast, the claim that

being in interaction with others is *per se* rewarding has never been put to the test, which most likely is due to the absence of appropriate methods.

We sought to explicitly assess whether experiencing engagement in social interaction recruits the reward system of the brain. In order to allow participants to engage in social interaction face-to-face and in real-time while they were in an MRI scanner, we used a within-subject adaption of the non-verbal Turing test introduced in ***study 3***. In this version, each interaction comprised five trials in which the virtual agent would either engage in joint or non-joint attention. The experiment consisted of two phases in which the human interaction partner was either introduced as naïve to participants' task or as explicitly cooperative. This permitted an exploration of the neural integration of another individual's gaze reactions during naïve and cooperative interaction. More importantly, however, this enabled investigating whether social interaction *per se* is sufficient to recruit reward-related neurocircuitry, or whether the introduction of a cooperative context is necessary. A major advantage of this setup the fact that the experience of engagement in social interaction is not defined *a priori* as an independent variable as in other studies employing human-computer distinctions, but emerges through the interaction dynamics.

Results show that the subjective experience of human interaction is correlated with activation of the ventral striatum (VS) and mOFC (Figure 6A), while experiencing non-social interaction is associated with increased neural activity in a fronto-parietal network related to attentional processing (Figure 6B). Behavioral data revealed that the time course of the integration of information during social interaction differs between naïve and cooperative interaction. During naïve interaction, already the first two trials (i.e. whether they are JA or NJA trials) exerted significant influence on the humanness judgment. On the contrary, during cooperation the influence of trial type increased towards the end of an interaction block. Based on these behavioral findings, additional fMRI analyses demonstrated that experiencing the interaction with a naïve and a cooperative interaction partner differentially modulates striatal activity. During naïve interaction, the striatum signals an early and putatively automatic preference subserving mechanisms of impression formation. During cooperation, it encodes the accumulation of value – i.e. the buildup of evidence that the interaction partner is actually human. Analyses based on the independent variable (i.e. the manipulation of the agent's gaze behavior) show that in the naïve condition

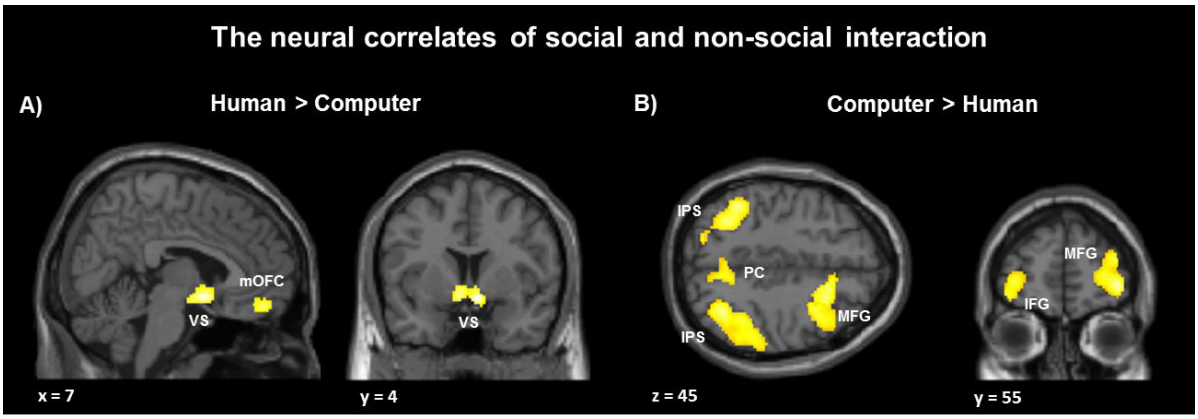

Figure 6. A) Neural correlates of social and non-social interaction. Greater activation of VS and mOFC during the experience of social interaction. B) Greater activation across the FPAN (including IPS, PC, MFG, IFG) during the experience of non-social interaction.

single events of joint attention recruit the social brain network irrespective of the condition without any integration of gaze reactions over the time course of an interaction block. In contrast, single events are stripped of their social salience during cooperation. The cooperative instruction appears to introduce an overarching mindset in which only contingency matters: Positive contingency is predominantly encoded by the dorsal striatum whereas negative contingency recruits the FPAN. The interplay of these two systems suggests a distribution of cognitive resources which might be required to ensure that joint attention – a social cue which is naturally meaningful and rewarding – is only associated with salience when the *overall* congruency of reactions is high and thereby indicative of a cooperative interaction partner.

These results provide first-time evidence that the subjective experience of social interaction is sufficient to recruit the reward system *even without* the presence of an explicitly cooperative interactor and, therefore, provide long sought confirmation of a neural motivation for social interaction. Moreover, the context of an interaction – that is, whether we interact with a naïve or a cooperative interaction partner – does not only influence our behavioral evaluation of her reactions to our own actions, but also has a great impact on the neural integration of these reactions.

6 General Discussion

The studies in the present thesis attempted to contribute to an understanding of the cognitive mechanisms and neural processes contributing to online social interaction. The development of an interactive eye-tracking paradigm (***study 1***) enabled us to study temporal and dynamic features of self-initiated processes of social gaze, such as gaze-following and joint attention, thereby opening a window to an understanding of gaze dynamics in triadic social interactions (***study 2***). Moreover, obtaining these data on a single-trial level made it possible to use sequences of gaze trials to construct basic and realistic interactions serving as the building blocks of the 'non-verbal Turing test' (***study 3***). Although ***study 3*** represented a somewhat explorative endeavor into understanding the mind in interaction, results demonstrated that the signals we evaluate as indicative of real human interaction crucially rely on the context of the interaction, i.e. whether someone is explicitly cooperating with us or not. Finally, in ***study 4*** an fMRI adaptation of this paradigm demonstrated that also the neural evaluation of these signals differs depending on whether we interact with a cooperator or not. Most intriguingly, however, results showed that explicit cooperation is not necessary to recruit the reward system during online social interaction but rather that the mere subjective experience of being in interaction is intrinsically rewarded as demonstrated by the activation of the ventral striatum and the mOFC. Importantly, social interaction was mediated by virtual agents instead of live faces. These results challenge the previous findings indicating that live faces are required to trigger mentalizing processes (Pönkänen et al., 2008; Wheatley et al., 2011), and thus strongly encourage the application of anthropomorphic virtual agents in the study of social interaction.

6.1 Motivational Aspects of Social Interaction

First speculations for the involvement of reward-related processes are supported by the findings of ***study 3***. This study focused on the expectations regarding the behavior of an alleged human interactor depending on whether this interactor was perceived as naïve, cooperative, or competitive. The main finding was that we are naturally predisposed to congruent gaze reactions (i.e. JA) as an indicator of an agent's humanness, while we

discount this expectation of mere congruency and accept generally contingent behavior (i.e. maximal JA *or* maximal NJA) as a cue to humanness in cases of explicit cooperation. This demonstrates that we are not blind to the actual contingency of gaze behavior but need an obvious reason to accept negatively contingent behavior as indicative of being in interaction with another human. This is likely to be related to the fact that gaze aversion is generally associated with negative emotional valence (e.g. Hietanen et al., 2008; Wirth et al., 2010). In comparison, JA is experienced as pleasant and recruits the VS as a function of the experienced pleasantness. This has been argued to represent an intrinsic neural motivation to initiate social interaction with another person (Schilbach et al., 2010). However, the study by Schilbach et al. (2010) does not provide unequivocal evidence for this claim. Participants in this study were informed that the other was explicitly instructed to either follow their gaze or avert its gaze in a given interaction block. In addition, they were informed about the other's instruction prior to each interaction block. The observed striatal activation might therefore be related to the feeling of being in control or to the confirmation of the expectation of another person's gaze-following behavior rather than to the experience of JA in itself (Delgado, 2007). This is supported by two studies which used a live face-to-face interaction setup to disentangle the neural correlates of initiating and responding to JA and failed to report differential activation of the VS for self-initiated JA (Redcay et al., 2010, 2012). These inconsistencies warranted a more direct investigation of the question whether the experience of online social interaction recruits the reward system. Indeed, previous neuroimaging studies suggested the recruitment of the reward system during social interactions. However, these studies comprised high-level neuroeconomic social interactions and operationalized human-computer distinctions explicitly as an independent variable (» Chapter 2.2.3). In contrast, the design of the Turing test required participants to determine the nature of their interaction partner in basic social interactions via the dynamics of the interaction process itself. The within-subject design of ***study 4*** provided the additional asset of allowing a comparison of neural activity during naïve and cooperative social interaction. This allowed answering the question whether *explicit cooperation* is necessary to turn social interaction into a rewarding experience. As ***study 4*** demonstrated, this is not case, which clearly demonstrates that the mere experience of active engagement in interaction with another person is intrinsically rewarding.

The findings of ***studies 3 and 4*** are consistent with proposals of an internal drive to

engage in interaction from early childhood on. Developmental studies have shown that behavioral contingencies inherent to social interactions between few-month old infants and their caregivers are experienced as joyful and pleasant (e.g. Rochat, 2001). Similarly, in the double television experiments by Murray and Trevarthen, babies interacted with their mothers via a television screen in real-time or saw recordings of their mothers' previous interaction behavior, which disrupted the behavioral contingency and caused severe emotional distress (Murray and Trevarthen, 1985). This has lend support to Trevarthen's concept of 'primary intersubjectivity' which suggests that from the very beginning of life there is a sharing of experience by elementary reciprocal interactions which are described as enjoyable both for infant and adult (Trevarthen, 1979). Such early interactions – so-called protoconversations – are devoid of any references to the outside world and therefore are not 'about' anything. Tomasello and colleagues noted that they require not only the capacity, but also the motivation to share emotions in the form of such simple interactions (Tomasello et al., 2005). At later stages of development, this motivation continually expands to sharing perceptions, goals, and activities, as measured by children's active pursuit of an engagement with others in order to share. Although Tomasello never claims explicitly that social interaction is rewarded by the brain, this is an implicit consequence of his postulate of an innate 'motivation to share' (Tomasello et al., 2005; Tomasello and Carpenter, 2007; Tomasello, 2009). As a final notion, the 'social brain hypothesis' shall be mentioned here, which argues that the complexity and extraordinary size of the human brain has emerged due to the evolutionary pressures created by living in groups (Dunbar, 1998). Specifically, he has argued that the limiting factor of group size is the information-processing capacity of the brain. Obviously, the question of whether intelligence enabled social cognition or whether social cognition enabled intelligence is a "chicken-and-egg question" (Adolphs, 2009, p. 699) which cannot be answered here. Nonetheless, it is conceivable that an intrinsic motivation to engage in social interaction with conspecifics might have co-evolved as a selection advantage. Taken together, however, it must be noted that despite providing first-time evidence for the rewarding value of social encounters, the data of ***studies 3 and 4*** can only serve as starting point for research on the motivational drive to interact. Further research is needed to investigate whether this is a central aspect of social interactions and which discrete factors of social interaction modulate the reward system.

6.2 Investigating Social Motivation in Autism

Future research might help to understand impairments of social cognition in clinical populations. While it has been argued throughout this thesis that social interaction is rewarding, this appears not to be true in autism spectrum disorders (ASD). ASD is characterized by repetitive, stereotyped behaviors, deficits in communication, and a severe impairment of social interaction (Baron-Cohen and Belmonte, 2005). Moreover, this impairment does not only seem to include the process of interaction but also the motivation to interact – that is, autistic persons do not seem to *want* to interact with others (Kohls et al., 2012). This 'social anhedonia' has been demonstrated across a wide range of social phenomena. For example, children with ASD prefer non-social to social stimuli and are less likely to help others spontaneously or to engage in shared collaborative activities. Furthermore, persons with ASD often report not having friends without feeling lonely, avoid eye contact, and do not initiate joint attention with others. These observations have recently led to the proposal of a 'social motivation theory of autism' which suggests deficits in the neural mechanisms supporting social rewards (Chevallier et al., 2012). Reward processing is usually divided into two components – an anticipatory phase of 'wanting' and a phase of reward 'liking'. One of the main claims of the social motivation hypothesis is that rather the 'wanting' than the 'liking' component is disrupted in ASD. At the moment, neuroimaging studies on the processing of social rewards are sparse. Interestingly, however, 'wanting' is mainly associated with activity of the VS (Peters and Büchel, 2010), which has been shown to be activated during the experience of social interaction in ***study 4***. It is thus conceivable that the non-verbal Turing test provides a suitable tool to scrutinize the neural mechanisms underlying the deficits in social motivation observed in ASD. Understanding the differences in reward-related activity during social encounters might be crucial for an understanding of both the 'healthy' and the autistic brain. Lastly, uncovering social motivational deficits of ASD is of great importance for the development of suitable therapy programs.

6.3 From One Mind to Two Minds

A further expansion of the present studies would consist of the implementation of dual eye-tracking setups which involve two individuals engaging in gaze-based interactions (Carletta et al., 2010). Very recent work by our group resulted in the development of such a system

which enables two individuals to engage in gaze-based interaction with each other while their gaze behavior is visualized by virtual agents. For a quantification of gaze parameters beyond gaze direction, this platform allows very fine-grained behavioral measurements and allows describing interactive gaze behavior in terms of direction, scan path length and area, number of saccades and fixations, fixation duration, and saccade velocity. In comparison to the interactive paradigm presented in this thesis and live video-streams, this platform has the major advantage that true real-time interactions between two individuals can be combined with the benefits of virtual reality methods (Barisic et al., *in press*). The avatars used to display participants' gaze behavior can be controlled for facial appearance and expressions whilst the microstructure of interactive gaze behavior is maintained. Furthermore, the environment surrounding the avatars can be modulated in a gaze-contingent fashion, thereby enabling a plethora of real-life as well as game-like social interactions in a highly realistic but controlled fashion. Very recently, live video-streams have been used in hyperscanning experiments on social gaze (Saito et al., 2010; Tanabe et al., 2012). Although these studies provided a proof-of-principle for the applicability of this method rather than novel insights into interactive gaze behavior, the possibility of combining the dual eye-tracking setup with hyperscanning has enormous potential for understanding the neural correlates of gaze-based interaction online.

7 Conclusion

The work reported in this thesis is altogether of an explorative nature. Not surprisingly, this opens numerous new questions to be addressed in future studies. Two major challenges can be identified at the present stage. Firstly, in the introduction of the second-person approach to other minds (» Chapter 2.2.1.) it has been noted that *offline* and *online* social cognition should differ from one another with respect to the underlying cognitive processes and neural mechanisms. Although the studies comprising this thesis offer valuable information about online mechanisms, they do not allow a direct comparison of *being in interaction* and *observing interaction*. Furthermore, such a comparison using the same kind of paradigm is of paramount importance for the viability of a second-person approach to social cognition (Przyrembel et al., 2012). A potential starting point for addressing this question might therefore be to design an observational version of the non-verbal Turing test in which two individuals engage in the task while being observed by another participant. The observer would then be asked to judge whether the participant who is initiating the gaze trials rates a given interaction sequence as social or non-social. The comparison of humanness ratings in the offline and the online version might provide a first measure of whether our experience of an interaction is changed by our active participation in this interaction. Secondly, the discovery that the subjective experience of social interaction is sufficient to recruit the reward system raises questions about the function of reward processes during cooperative activities. Are we born to cooperate, as suggested by Tomasello (2009), or are we born to interact, as the data of study 4 suggest? It might be entirely possible that we cooperate due to a much more basal motive – namely to sustain social interactions. As cooperation usually entails the engagement in social interactions, future research should focus on investigating commonalities as well as differences in the activation of the reward system during unconstrained and cooperative interactions.

8 References

Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annu. Rev. Psychol.* 60, 693–716.

Adolphs, R. (2010). What does the amygdala contribute to social cognition? *Ann. N. Y. Acad. Sci.* 1191, 42–61.

Adolphs, R., Tranel, D., Damasio, H., and Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* 372, 669–672.

Allison, Puce, and McCarthy (2000). Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* 4, 267–278.

Amodio, D. M., and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277.

Argyle, M., Ingham, R., Alkema, F., and McCallin, M. (1973). The Different Functions of Gaze. *Semiotica* 7, 19–32.

Auvray, M., Lenay, C., and Stewart, J. (2009). Perceptual interactions in a minimalist virtual environment. *New Ideas Psychol.* 27, 32–47.

Babiloni, F., Cincotti, F., Mattia, D., Mattiocco, M., De Vico Fallani, F., Tocci, A., Bianchi, L., Marciani, M. G., and Astolfi, L. (2006). Hypermethods for EEG hyperscanning. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 1, 3666–3669.

Baess, P., Zhdanov, A., Hirvenkari, L., Mäkelä, J. P., Jousmäki, V., and Hari, R. (2012). MEG dual scanning: a procedure to study real-time auditory interaction between two persons. *Front. Neurosci.* 6, 83.

Bailenson, J. N., Beall, A. C., Loomis, J., Blascovich, J., and Turk, M. (2005). Transformed Social Interaction, Augmented Gaze, and Social Influence in Immersive Virtual Environments. *Hum. Commun. Res.* 31, 511–537.

Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2003). Interpersonal Distance in Immersive Virtual Environments. *Pers. Soc. Psychol. Bull.* 29, 819–833.

Barisic, I., Timmermans, B., Pfeiffer, U. J., Vogeley, K., and Schilbach, L. (*in press*). Using dual eye-tracking to investigate real-time social interactions. in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (Paris).

Baron-Cohen, S., and Belmonte, M. K. (2005). Autism: a window onto the development of the social and the analytic brain. *Annu. Rev. Neurosci.* 28, 109–126.

Baron-Cohen, S., Leslie, A. M., and Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition* 21, 37–46.

Baumeister, R. F., and Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* 117, 497–529.

Becchio, C., Sartori, L., and Castiello, U. (2010). Toward You: The Social Side of Actions. *Curr. Dir. Psychol. Sci.* 19, 183–188.

Bente, G., Eschenburg, F., and Krämer, N. C. (2007). "Virtual Gaze. A Pilot Study on the Effects of Computer Simulated Gaze in Avatar-Based Conversations," in *Virtual Reality*, ed. R. Shumaker (Berlin, Heidelberg: Springer), 185–194.

Berry, D. S., and McArthur, L. Z. (1986). Perceiving character in faces: The impact of age-related craniofacial changes on social perception. *Psychol. Bull.* 100, 3–18.

Bohil, C. J., Alicea, B., and Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nat. Rev. Neurosci.* 12, 752–762.

Brass, M., Schmitt, R. M., Spengler, S., and Gergely, G. (2007). Investigating action understanding: inferential processes versus action simulation. *Curr. Biol.* 17, 2117–2121.

Bristow, D., Rees, G., and Frith, C. D. (2007). Social interaction modifies neural response to gaze shifts. *Soc. Cogn. Affect. Neurosci.* 2, 52–61.

Brown, E. C., and Brüne, M. (2012). The role of prediction in social neuroscience. *Front. Neurosci.* 6, 147.

Buber, M. (2004). *I and Thou*. London, UK: Continuum International Publishing Group.

Carletta, J., Hill, R. L., Nicol, C., Taylor, T., De Ruiter, J. P., and Bard, E. G. (2010). Eyetracking for two-person tasks with manipulation of a virtual world. *Behav. Res. Methods* 42, 254–265.

Carlin, J. D., and Calder, A. J. (*in press*). The neural basis of eye gaze processing. *Curr. Opin. Neurobiol.*

Carlin, J. D., Calder, A. J., Kriegeskorte, N., Nili, H., and Rowe, J. B. (2011). A head view-invariant representation of gaze direction in anterior superior temporal sulcus. *Curr. Biol.* 21, 1817–1821.

Carlin, J. D., Rowe, J. B., Kriegeskorte, N., Thompson, R., and Calder, A. J. (2012). Direction-sensitive codes for observed head turns in human superior temporal sulcus. *Cereb. Cortex* 22, 735–744.

Carpenter, M., Nagell, K., and Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* 63, i–vi, 1–143.

Carrington, S. J., and Bailey, A. J. (2009). Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Hum. Brain Mapp.* 30, 2313–2335.

Cary, M. S. (1978). The Role of Gaze in the Initiation of Conversation. *Soc. Psychol.* 41, 269.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., and Schultz, R. T. (2012). The social motivation theory of autism. *Trends Cogn. Sci.* 16, 231–239.

Clark, H. H. (1996). *Using Language*. 2nd ed. Cambridge, UK: Cambridge University Press.

Cloninger, C. R., Svrakic, D. M., and Przybeck, T. R. (1993). A psychobiological model of temperament and character. *Arch. Gen. Psychiatry* 50, 975–990.

Conty, L., and Grèzes, J. (2012). Look at me, I'll remember you. *Hum. Brain Mapp.* 33, 2428–2440.

Csibra, G., and Gergely, G. (2009). Natural pedagogy. *Trends Cogn. Sci.* 13, 148–153.

Cunningham, W. A., Johnson, M. K., Raye, C. L., Chris Gatenby, J., Gore, J. C., and Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychol. Sci.* 15, 806–813.

De Jaegher, H. D., and Di Paolo, E. D. (2007). Participatory sense-making. *Phenom. Cogn. Sci.* 6, 485–507.

De Jaegher, H., Di Paolo, E., and Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends Cogn. Sci.* 14, 441–447.

De Lange, F. P., Spronk, M., Willems, R. M., Toni, I., and Bekkering, H. (2008). Complementary systems for understanding action intentions. *Curr. Biol.* 18, 454–457.

Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T., and Meltzoff, A. N. (2004). The neural bases of cooperation and competition: an fMRI investigation. *Neuroimage* 23, 744–751.

Delgado, M. R. (2007). Reward-Related Responses in the Human Striatum. *Ann. N. Y. Acad. Sci.* 1104, 70–88.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain Res.* 91, 176–180.

Duchowski, A. T., Cournia, N., and Murphy, H. (2004). Gaze-Contingent Displays: A Review. *Cyberpsychol. Behav.* 7, 621–634.

Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., and Garnero, L. (2010). Inter-brain synchronization during social interaction. *PLoS ONE* 5, e12166.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evol. Anthr.* 6, 178–190.

Ekman, P., and Oster, H. (1979). Facial Expressions of Emotion. *Annu. Rev. Psychol.* 30, 527–554.

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neurosci. Biobehav. Rev.* 24, 581–604.

Farroni, T., Csibra, G., Simion, F., and Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9602–9605.

Fitzgerald, D. A., Angstadt, M., Jelsone, L. M., Nathan, P. J., and Phan, K. L. (2006). Beyond threat: amygdala reactivity across multiple expressions of facial affect. *Neuroimage* 30, 1441–1448.

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., and Frith, C. D. (1995). Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. *Cognition* 57, 109–128.

Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E., and Falk, A. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science* 318, 1305–1308.

Fox, J., Arena, D., and Bailenson, J. N. (2009). Virtual Reality. *J. Media Psychol.* 21, 95–113.

Frith, C. D. (2007). The social brain? *Phil. Trans. R. Soc. B* 362, 671–678.

Frith, C. D., and Frith, U. (2006). The neural basis of mentalizing. *Neuron* 50, 531–534.

Frith, U., and Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Phil. Trans. R. Soc. B* 358, 459–473.

Froese, T., and Fuchs, T. (2012). The extended body: a case study in the neurophenomenology of social interaction. *Phenom. Cogn. Sci.* 11, 205–235.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., and Frith, C. D. (2000). Reading the mind in cartoons and stories: an fMRI study of "theory of mind" in verbal and nonverbal tasks. *Neuropsychologia* 38, 11–21.

Gallagher, H. L., Jack, A. I., Roepstorff, A., and Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage* 16, 814–821.

Gazzola, V., Rizzolatti, G., Wicker, B., and Keysers, C. (2007). The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage* 35, 1674–1684.

Gillath, O., McCall, C., Shaver, P. R., and Blascovich, J. (2008). What Can Virtual Reality Teach Us About Prosocial Tendencies in Real and Virtual Environments? *Media Psychol.* 11, 259–282.

Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., and Haxby, J. V. (2007). Two takes on the social brain: a comparison of theory of mind tasks. *J. Cogn. Neurosci.* 19, 1803–1814.

Grant, E. C. (1969). Human Facial Expression. *Man* 4, 525–692.

Grynszpan, O., Nadel, J., Martin, J.-C., Simonin, J., Bailleul, P., Wang, Y., Gepner, D., Le Barillier, F., and Constant, J. (2011). Self-Monitoring of Gaze in High Functioning Autism. *J. Autism Dev. Disord.*

Grynszpan, O., Simonin, J., Martin, J.-C., and Nadel, J. (2012). Investigating social gaze as an action-perception online performance. *Front. Neurosci.* 6, 94.

Haith, M. M., Bergman, T., and Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science* 198, 853–855.

Han, S., Jiang, Y., Humphreys, G. W., Zhou, T., and Cai, P. (2005). Distinct neural substrates for the perception of real and virtual visual worlds. *NeuroImage* 24, 928–935.

Happé, F., Ehlers, S., Fletcher, P., Frith, U., Johansson, M., Gillberg, C., Dolan, R., Frackowiak, R., and Frith, C. (1996). "Theory of mind" in the brain. Evidence from a PET scan study of Asperger syndrome. *Neuroreport* 8, 197–201.

Hardin, C. D., and Higgins, E. T. (1996). "Shared Reality: How social verification makes the subjective objective," in *Handbook of motivation and cognition: The interpersonal context*, eds. E. T. Higgins and R. M. Sorrentino (New York: Guilford).

Hartwright, C. E., Apperly, I. A., and Hansen, P. C. (2012). Multiple roles for executive control in belief-desire reasoning: distinct neural networks are recruited for self perspective inhibition and complexity of reasoning. *Neuroimage* 61, 921–930.

Hein, G., and Knight, R. T. (2008). Superior temporal sulcus--It's my area: or is it? *J. Cogn. Neurosci.* 20, 2125–2136.

Hietanen, J. K., Leppänen, J. M., Peltola, M. J., Linna-aho, K., and Ruuhiala, H. J. (2008). Seeing direct and averted gaze activates the approach-avoidance motivational brain systems. *Neuropsychologia* 46, 2423–2430.

Hobson, R. P. (1991). Against the theory of "Theory of Mind". *Br. J. Dev. Psychol.* 9, 33–51.

Hoyt, C. L., Blascovich, J., and Swinth, K. R. (2003). Social Inhibition in Immersive Virtual Environments. *Presence-Teleop. Virt.* 12, 183–195.

Itier, R. J., and Batty, M. (2009). Neural bases of eye and gaze processing: the core of social cognition. *Neurosci. Biobehav. Rev.* 33, 843–863.

Kätsyri, J., Hari, R., Ravaja, N., and Nummenmaa, L. (2012). The Opponent Matters: Elevated fMRI Reward Responses to Winning Against a Human Versus a Computer Opponent During Interactive Video Game Playing. *Cereb. Cortex*.

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., and Montague, P. R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83.

Kircher, T., Blümel, I., Marjoram, D., Lataster, T., Krabbendam, L., Weber, J., Van Os, J., and Krach, S. (2009). Online mentalising investigated with functional MRI. *Neurosci. Lett.* 454, 176–181.

Kirk, U., Downar, J., and Montague, P. R. (2011). Interoception drives increased rational decision-making in meditators playing the ultimatum game. *Front. Neurosci.* 5, 49.

Kobayashi, H., and Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *J. Hum. Evol.* 40, 419–435.

Kohls, G., Chevallier, C., Troiani, V., and Schultz, R. T. (2012). Social "wanting" dysfunction in autism: neurobiological underpinnings and treatment implications. *J. Neurodev. Dis.* 4, 10.

Konvalinka, I., and Roepstorff, A. (2012). The two-brain approach: how can mutually interacting brains teach us something about social interaction? *Front. Neurosci.* 6, 215.

Krach, S., Paulus, F. M., Bodden, M., and Kircher, T. (2010). The rewarding nature of social interactions. *Front. Neurosci.* 4, 22.

Kuzmanovic, B., Georgescu, A. L., Eickhoff, S. B., Shah, N. J., Bente, G., Fink, G. R., and Vogeley, K. (2009). Duration matters: dissociating neural correlates of detection and evaluation of social gaze. *Neuroimage* 46, 1154–1163.

Lahnakoski, J. M., Glerean, E., Salmi, J., Jääskeläinen, I. P., Sams, M., Hari, R., and Nummenmaa, L. (2012). Naturalistic FMRI mapping reveals superior temporal sulcus as the hub for the distributed brain network for social perception. *Front. Neurosci.* 6, 233.

Lebreton, M., Barnes, A., Miettunen, J., Peltonen, L., Ridler, K., Veijola, J., Tanskanen, P., Suckling, J., Jarvelin, M.-R., Jones, P. B., et al. (2009). The brain structural disposition to social interaction. *Eur. J. Neurosci.* 29, 2247–2252.

LeDoux, J. E. (2000). Emotion Circuits in the Brain. *Annu. Rev. Neurosci.* 23, 155–184.

Liepelt, R., Von Cramon, D. Y., and Brass, M. (2008). How do we infer others' goals from non-stereotypic actions? The outcome of context-sensitive inferential processing in right inferior parietal and posterior temporal cortex. *Neuroimage* 43, 784–792.

Liu, X., Powell, D. K., Wang, H., Gold, B. T., Corbly, C. R., and Joseph, J. E. (2007). Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. *J. Neurosci.* 27, 4587–4597.

Mar, R. A., Kelley, W. M., Heatherton, T. F., and Macrae, C. N. (2007). Detecting agency from the biological motion of veridical vs animated agents. *Soc. Cogn. Affect. Neurosci.* 2, 199–205.

Marsh, K. L., Richardson, M. J., and Schmidt, R. C. (2009). Social Connection Through Joint Action and Interpersonal Coordination. *Topics Cog. Sci.* 1, 320–339.

Mason, M. F., Tatkow, E. P., and Macrae, C. N. (2005). The look of love: gaze shifts and person perception. *Psychol. Sci.* 16, 236–239.

Materna, S., Dicke, P. W., and Thier, P. (2008). Dissociable roles of the superior temporal sulcus and the intraparietal sulcus in joint attention: a functional magnetic resonance imaging study. *J. Cogn. Neurosci.* 20, 108–119.

Mitchell, J. P., Macrae, C. N., and Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron* 50, 655–663.

Moll, H., and Tomasello, M. (2007). Cooperation and human cognition: the Vygotskian intelligence hypothesis. *Phil. Trans. R. Soc. B* 362, 639–648.

Montague, P. R., Berns, G. S., Cohen, J. D., McClure, S. M., Pagnoni, G., Dhamala, M., Wiest, M. C., Karpov, I., King, R. D., Apple, N., et al. (2002). Hyperscanning: simultaneous fMRI during linked social interactions. *Neuroimage* 16, 1159–1164.

Mori, M. (1970). The Uncanny Valley. *Energy* 7, 33–35.

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., and Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature* 383, 812–815.

Mundy, P., and Newell, L. (2007). Attention, Joint Attention, and Social Cognition. *Curr. Dir. Psychol. Sci.* 16, 269–274.

Murray, L., and Trevarthen, C. (1985). "Emotional regulation of interaction between two-month-olds and their mothers," in *Social perception in infants*, eds. T. Field and N. Fox (New Jersey: Ablex Publishing), 101–125.

Niedenthal, P. M. (2007). Embodying Emotion. *Science* 316, 1002–1005.

Nummenmaa, L., and Calder, A. J. (2009). Neural mechanisms of social attention. *Trends Cogn. Sci.* 13, 135–143.

Pelphrey, K. A., Morris, J. P., and McCarthy, G. (2004a). Grasping the Intentions of Others: The Perceived Intentionality of an Action Influences Activity in the Superior Temporal Sulcus during Social Perception. *J. Cogn. Neurosci.* 16, 1706–1716.

Pelphrey, K. A., Singerman, J. D., Allison, T., and McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: the influence of context. *Neuropsychologia* 41, 156–170.

Pelphrey, K. A., Viola, R. J., and McCarthy, G. (2004b). When Strangers Pass Processing of Mutual and Averted Social Gaze in the Superior Temporal Sulcus. *Psychol. Sci.* 15, 598–603.

Peters, J., and Büchel, C. (2010). Neural representations of subjective reward value. *Behav. Brain Res.* 213, 135–141.

Pfeiffer, U., and Weidner, R. (2013). "Augenbewegungen", in *Funktionelle MRT in Psychiatrie und Neurologie*, eds. F. Schneider and G. R. Fink (Heidelberg: Springer).

Pönkänen, L. M., Hietanen, J. K., Peltola, M. J., Kauppinen, P. K., Haapalainen, A., and Leppänen, J. M. (2008). Facing a real person: an event-related potential study. *NeuroReport* 19, 497–501.

Pönkänen, L. M., Peltola, M. J., and Hietanen, J. K. (2011). The observer observed: frontal EEG asymmetry and autonomic responses differentiate between another person's direct and averted gaze when the face is seen live. *Int. J. Psychophysiol.* 82, 180–187.

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526.

Przyrembel, M., Smallwood, J., and Singer, T. (2012). Illuminating the dark matter of social neuroscience: Considering the problem of social interaction from philosophical, psychological, and neuroscientific perspectives. *Front. Neurosci.* 6, 190.

Rapson, R. L., Hatfield, E., and Cacioppo, J. T. (1994). *Emotional Contagion*. Cambridge, UK: Cambridge University Press.

Redcay, E. (2008). The superior temporal sulcus performs a common function for social and speech perception: implications for the emergence of autism. *Neurosci. Biobehav. Rev.* 32, 123–142.

Redcay, E., Dodell-Feder, D., Pearrow, M. J., Mavros, P. L., Kleiner, M., Gabrieli, J. D. E., and Saxe, R. (2010). Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. *Neuroimage* 50, 1639–1647.

Redcay, E., Kleiner, M., and Saxe, R. (2012). Look at this: the neural correlates of initiating and responding to bids for joint attention. *Front. Neurosci.* 6, 169.

Reddy, V., and Morris, P. (2004). Participants Don't Need Theories Knowing Minds in Engagement. *Theory Psychology* 14, 647–665.

Riley, M. A., Richardson, M. J., Shockley, K., and Ramenzoni, V. C. (2011). Interpersonal Synergies. *Front. Psychol.* 2.

Rilling, J. K., and Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annu. Rev. Psychol.* 62, 23–48.

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage* 22, 1694–1703.

Rizzolatti, G., and Craighero, L. (2004). The Mirror-Neuron System. *Annu. Rev. Neurosci.* 27, 169–192.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., and Fazio, F. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Exp. Brain Res.* 111, 246–252.

Rizzolatti, G., and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat. Rev. Neurosci.* 11, 264–274.

Rochat, M. J., Caruana, F., Jezzini, A., Escola, L., Intskirveli, I., Grammont, F., Gallese, V., Rizzolatti, G., and Umiltà, M. A. (2010). Responses of mirror neurons in area F5 to hand and tool grasping observation. *Exp. Brain Res.* 204, 605–616.

Rochat, P. R. (2001). Social contingency detection and infant development. *Bull. Menninger Clin.* 65, 347–360.

Saito, D. N., Tanabe, H. C., Izuma, K., Hayashi, M. J., Morito, Y., Komeda, H., Uchiyama, H., Kosaka, H., Okazawa, H., Fujibayashi, Y., et al. (2010). “Stay tuned”: inter-individual neural synchronization during mutual gaze and joint attention. *Front. Neurosci.* 4, 127.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2003). The Neural Basis of Economic Decision-Making in the Ultimatum Game. *Science* 300, 1755–1758.

Saxe, R. (2006). Uniquely human social cognition. *Curr. Opin. Neurobiol.* 16, 235–239.

Saxe, R., and Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage* 19, 1835–1842.

Saxe, R. R., Whitfield-Gabrieli, S., Scholz, J., and Pelphrey, K. A. (2009). Brain regions for perceiving and reasoning about other people in school-aged children. *Child Dev.* 80, 1197–1209.

Saxe, R., and Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia* 43, 1391–1399.

Schank, R. C., and Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Oxford, England: Lawrence Erlbaum.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., and Vogeley, K. (*in press*). Toward a second-person neuroscience. *Behav. Brain Sci.*

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., Shah, N. J., Fink, G. R., and Vogeley, K. (2010). Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. *J. Cogn. Neurosci.* 22, 2702–2715.

Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., and Vogeley, K. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia* 44, 718–730.

Schultz, R. T., Grelotti, D. J., Klin, A., Kleinman, J., Van der Gaag, C., Marois, R., and Skudlarski, P. (2003). The role of the fusiform face area in social cognition: implications for the pathobiology of autism. *Phil. Trans. R. Soc. B* 358, 415–427.

Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends Cogn. Sci.* 10, 70–76.

Senju, A., and Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends Cogn. Sci.* 13, 127–134.

Shamay-Tsoory, S. G., and Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia* 45, 3054–3067.

Shepherd, S. V. (2010). Following gaze: gaze-following behavior as a window into social cognition. *Front. Neurosci.* 4, 5.

Strick, M., Holland, R. W., and Van Knippenberg, A. (2008). Seductive eyes: attractiveness and direct gaze increase desire for associated objects. *Cognition* 106, 1487–1496.

Tamir, D. I., and Mitchell, J. P. (2010). Neural correlates of anchoring-and-adjustment during mentalizing. *Proc. Natl. Acad. Sci. U.S.A* 107, 10827–10832.

Tanabe, H. C., Kosaka, H., Saito, D. N., Koike, T., Hayashi, M. J., Izuma, K., Komeda, H., Ishitobi, M., Omori, M., Munesue, T., et al. (2012). Hard to "tune in": neural mechanisms of live face-to-face interaction with high-functioning autistic spectrum disorder. *Front. Neurosci.* 6, 268.

Tognoli, E., Lagarde, J., DeGuzman, G. C., and Kelso, J. A. S. (2007). The phi complex as a neuromarker of human social coordination. *Proc. Natl. Acad. Sci. U.S.A.* 104, 8190–8195.

Tomasello, M. (2009). *Why We Cooperate*. Cambridge, MA: The MIT Press.

Tomasello, M., and Carpenter, M. (2007). Shared intentionality. *Dev. Sci.* 10, 121–125.

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–691.

Tomlin, D., Kayali, M. A., King-Casas, B., Anen, C., Camerer, C. F., Quartz, S. R., and Montague, P. R. (2006). Agent-specific responses in the cingulate cortex during economic exchanges. *Science* 312, 1047–1050.

Trevarthen, C. (1979). "Communication and cooperation in early infancy. a description of primary intersubjectivity", in *Before Speech: The Beginning of Human*

Communication, eds. K. Scherer and P. Ekman (London, UK: Cambridge University Press), 321–347.

Turing, A. (1950). Computing Machinery and Intelligence. *Mind* LIX, 433–460.

Tylén, K., Allen, M., Hunter, B. K., and Roepstorff, A. (2012). Interaction vs. observation: distinctive modes of social cognition in human brain and behavior? A combined fMRI and eye-tracking study. *Front. Neurosci.* 6, 331.

Umiltà, M. A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., Jezzini, A., Gallese, V., and Rizzolatti, G. (2008). When pliers become fingers in the monkey motor system. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2209–2213.

Van Overwalle, F., and Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage* 48, 564–584.

Vogeley, K., and Bente, G. (2010). "Artificial humans": Psychology and neuroscience perspectives on embodiment and nonverbal communication. *Neural Netw.* 23, 1077–1090.

Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happé, F., Falkai, P., Maier, W., Shah, N. J., Fink, G. R., and Zilles, K. (2001). Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage* 14, 170–181.

Vogeley, K., and Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends Cogn. Sci.* 7, 38–42.

Walker-Smith, G. J., Gale, A. G., and Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception* 6, 313 – 326.

Walter, H., Adenzato, M., Ciaramidaro, A., Enrici, I., Pia, L., and Bara, B. G. (2004). Understanding intentions in social interaction: the role of the anterior paracingulate cortex. *J. Cogn. Neurosci.* 16, 1854–1863.

Wheatley, T., Weinberg, A., Looser, C., Moran, T., and Hajcak, G. (2011). Mind Perception: Real but Not Artificial Faces Sustain Neural Activity beyond the N170/VPP. *PLoS ONE* 6, e17960.

Williams, J. H. G., Waiter, G. D., Perra, O., Perrett, D. I., and Whiten, A. (2005). An fMRI study of joint attention experience. *NeuroImage* 25, 133–140.

Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G. R., and Vogeley, K. (2010). It's in your eyes--using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Soc. Cogn. Affect. Neurosci.* 5, 98–107.

Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128.

Wirth, J. H., Sacco, D. F., Hugenberg, K., and Williams, K. D. (2010). Eye gaze as relational evaluation: averted eye gaze leads to feelings of ostracism and relational devaluation. *Pers. Soc. Psychol. Bull.* 36, 869–882.

Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Phil. Trans. R. Soc. B* 358, 593–602.

Wolpert, D. M., Ghahramani, Z., and Flanagan, J. R. (2001). Perspectives and problems in motor learning. *Trends Cogn. Sci.* 5, 487–494.

Zaki, J., Hennigan, K., Weber, J., and Ochsner, K. N. (2010). Social cognitive conflict resolution: contributions of domain-general and domain-specific neural systems. *J. Neurosci.* 30, 8481–8488.

Appendix

Wilms M*, Schilbach L*, **Pfeiffer UJ**, Bente G, Fink GR, Vogeley K, 2010. It's in your eyes – using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Social Cognitive and Affective Neuroscience*, 5(1), pp. 98 – 107. (*equal contributions)

Pfeiffer UJ, Schilbach L, Jording M, Timmermans B, Bente G, Vogeley K, 2012. Eyes on the mind: investigating the influence of gaze dynamics on the perception of others in real-time social interaction. *Frontiers in Psychology*, 3, p. 537.

Pfeiffer UJ, Timmermans B, Bente G, Vogeley K, Schilbach L, 2011. A non-verbal Turing test: differentiating mind from machine in gaze-based social interaction. *PLoS One*, 6(11), e27591.

Pfeiffer UJ, Schilbach L, Timmermans B, Kuzmanovic B, Georgescu AL, Bente G, Vogeley K. (*submitted*). Why we interact: On the functional role of the ventral striatum during real-time social interaction.

doi:10.1093/scan/nsq024 SCAN (2010) 5, 98–107

Tools of the Trade

It's in your eyes—using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience

Marcus Wilms,[1,*] Leonhard Schilbach,[2,*] Ulrich Pfeiffer,[2] Gary Bente,[3] Gereon R. Fink,[1,4] and Kai Vogeley[2]
[1]Institute of Neurosciences and Medicine, Research Centre Juelich, Juelich, [2]Department of Psychiatry, [3]Department of Social Psychology, and [4]Department of Neurology, University of Cologne, Cologne, Germany

The field of social neuroscience has made remarkable progress in elucidating the neural mechanisms of social cognition. More recently, the need for new experimental approaches has been highlighted that allow studying social encounters in a truly interactive manner by establishing 'online' reciprocity in social interaction. In this article, we present a newly developed adaptation of a method which uses eyetracking data obtained from participants in real time to control visual stimulation during functional magnetic resonance imaging, thus, providing an innovative tool to generate gaze-contingent stimuli in spite of the constraints of this experimental setting. We review results of two paradigms employing this technique and demonstrate how gaze data can be used to animate a virtual character whose behavior becomes 'responsive' to being looked at allowing the participant to engage in 'online' interaction with this virtual other in real-time. Possible applications of this setup are discussed highlighting the potential of this development as a new 'tool of the trade' in social cognitive and affective neuroscience.

Keywords: magnetic resonance imaging-compatible interactive eyetracking; truly interactive paradigms; gaze feedback; social cognition

INTRODUCTION

Social neuroscience has helped to shed light upon the neural mechanisms underlying our ability to understand other minds under the headings of 'theory of mind' or 'mentalizing', commonly understood as the ability to represent other people's mental states (Frith and Frith, 2003, 2008). An increasing number of functional neuroimaging studies suggests that two large-scale neural networks are involved, namely the so-called mirror neuron system, comprising essentially the parietal and premotor cortices and the so-called 'social brain', comprising essentially the medial prefrontal, the temporopolar, the temporoparietal cortices and the amygdala (Keysers and Gazzola, 2007; Lieberman, 2007).

Most of these studies, however, rely on paradigms in which participants are asked to merely observe others ('offline' mentalizing; Schilbach *et al.*, 2006), while mentalizing during 'online' social interaction has only been studied by a minority of studies (e.g. Montague *et al.*, 2002; Eisenberger *et al.*, 2003), often making use of game theory paradigms from economics (e.g. Sanfey, 2007). Consequently, the need to develop 'interactive mind' paradigms that could provide a platform to systematically study the neural mechanisms of social interaction in an ecologically valid manner has been pointed out recently (Singer, 2006). 'Online' interaction crucially involves 'closing the loop' between interaction partners and establishing reciprocal relations where actions feed directly into the communication loop and elicit reactions which, in turn, may subsequently lead to reactions of the initiator and so forth. This has been referred to as adopting a 'second-person-perspective' (2PP; Reddy, 2003) which can be taken to suggest that awareness of mental states results from being psychologically engaged with someone and being an active participant of reciprocal interaction thereby establishing a subject-subject ('Me–You') rather than a subject–object ('Me–She/He') relationship. Paradigms that permit the systematic investigation of the reciprocity of interactions as well as the involvement of implicit and explicit processes will substantially enrich our knowledge of the neurobiology of social cognition (Frith and Frith, 2008).

The challenge for social neuroscience here will be twofold: a suitable experimental platform should allow real-time, 'online' interactions between participants and the social

Received 20 November 2009; Accepted 17 February 2010
Advance Access publication 11 March 2010

We are grateful to our colleagues from the MR group at the Institute for Neurosciences and Medicine, in particular Ralph Weidner and Sharam Mirzazade, for their assistance in acquiring eyetracking data inside the scanner. The study was supported by a grant of the German Ministery of Research and Education (BMBF) on 'social gaze' (G.B. and K.V.). L.S. was also supported by the Koeln Fortune Program of the Faculty of Medicine, University of Cologne.

Correspondence should be addressed to Leonhard Schilbach, Department of Psychiatry, University Hospital Cologne, Kerpener Str. 62, 50924 Cologne, Germany. E-mail: leonhard.schilbach@uk-koeln.de
*These authors contributed equally to this work.

stimuli while also providing means for experimental control over changes of the social stimuli. Here, we suggest that measurements of participants' gaze behavior inside the MR environment could be used to influence a virtual character's gaze behavior making it 'responsive' to the participant's gaze allowing to engage in 'online' interaction with the virtual other in real-time.

From a theoretical standpoint, gaze is known to be an important social cue in dyadic interaction indicating interest or disinterest, regulating intimacy levels, seeking feedback and expressing emotions (e.g. Argyle and Cook, 1976; Emery, 2000). In addition, gaze can also influence object perception (Becchio, *et al.*, 2008) by means of establishing triadic relations between two observers and an object onto which the interactors can look 'together' and thereby establish 'joint attention' (Moore and Dunham, 1995). From a methodological standpoint, gaze behavior represents one of the few ways in which participants can interact with stimuli naturally in spite of the movement constraints when lying inside an MR scanner. Gaze is a socially most salient nonverbal behavior, which can be reliably measured even within an functional magnetic resonance imaging (fMRI) setting and can, therefore, be used as input for interactive paradigms in neuroimaging studies.

This leads to the further question of how gaze input can be used to automatically control contingent behavior of a social stimulus to create the illusion of real-time interaction in such a setting. Pre-programmed and strictly controlled visual presentation of nonverbal behaviors in general and of gaze cues in particular, can be established by using anthropomorphic virtual characters. Such computer-generated characters have been suggested as a valuable tool for social neuroscience (e.g. Sanchez-Vives and Slater, 2005) as they convey social information to human observers (e.g. Bente *et al.*, 2001; Bente *et al.*, 2002, 2008a; Bailenson *et al.*, 2003) and cause reactions strikingly similar to those in real human interactions (e.g. Slater *et al.*, 2006). An increasing number of behavioral as well as fMRI studies has now used such stimuli to study different aspects of social cognition including gaze perception (e.g. Pelphrey *et al.*, 2005; Spiers *et al.*, 2006; Schilbach *et al.*, 2006; Bente *et al.*, 2007, 2008b; Park *et al.*, 2009; Kuzmanovic *et al.*, 2009).

To establish a paradigm that actively engages participants in the 2PP as opposed to being a 'detached' observer of social stimuli (from a 'third-person-perspective'; 3PP), we present a new technique that makes use of eyetracking data obtained from participants inside an MR scanner to control a virtual character's gaze behavior in real-time making it 'responsive' to the human observer's gaze (Schilbach *et al.*, in press). In this setup the eye movements of the participant become a means to 'probe' the behavior of the virtual other similar to real-life situations. Importantly, this also seems compatible with an 'enactive' account of social cognition which understands social cognition as bodily experiences resulting from an organism's adaptive actions upon salient and self-relevant aspects of the environment (e.g. Klin *et al.*, 2003) that feed back into the social interaction process.

Consequently, our setup promises to allow the exploration of the neural basis of processes of interpersonal alignment and the reciprocity inherent to social interaction, i.e. whether and how social cues are detected as contingent upon one's own behavior and how interaction partners initiate and respond to each other's actions (Schilbach *et al.*, in press). Both aspects seem highly relevant to make substantial progress in the field of social cognitive neuroscience and may lead to a reconsideration of the current emphasis on similarities between self- and other-related processing.

To implement these different eyetracking setups were tested in the fMRI setting to produce gaze-contingent stimuli. We review results of these different approaches which use eyetracking measurements overtly or covertly to drive MR-compatible experimental paradigms and underline the usability of this technique. Furthermore, we give examples for the applications of these interactive, eyetracking-based paradigms. Given the importance of gaze behavior during real-life social interaction, this approach, we suggest, provides a much needed, new 'tool of the trade' for the study of real-time 'online' interaction in social neuroscience.

METHODS

Interactive eyetracking setups

'Interactive eyetracking' relies on an MR-compatible eyetracking system that allows real-time data transmission to a visual stimulation controller. The controller receives the ongoing gaze data and adapts the visual stimulation according to preset task conditions and the volunteer's current gaze position on screen.

For stimulus delivery, different presentation devices were tested employing either a TFT screen or two different goggle systems. First, a custom-built, shielded TFT screen was used for the stimulus presentation at the rear end of the scanner ($14^\circ \times 8^\circ$ horizontal × vertical viewing angle, screen distance from volunteer's eyes: 245 cm). Volunteers watched the stimuli via a mirror mounted on the head coil. Volunteers' eye movements were monitored by means of an infrared camera (Resonance Technology, CA, USA). The camera and infra-red light source were mounted on the head coil using a custom-built gooseneck that allowed easy access to the volunteer's eyes without interfering with the visual stimulation (setup A). Second, stimuli were presented using MR-compatible goggles. Volunteers' eye movements were monitored by means of an infrared camera that was built into the goggles. In a 3T MR environment we used a VisuaStim^{TM} system ($30^\circ \times 22.5^\circ$ horizontal × vertical viewing angle; Resonance Technology, CA, USA; setup B_1) whereas in a 1.5 T environment we tested a Silent Vision^{TM} ($25.5^\circ \times 18^\circ$ horizontal × vertical viewing angle; AvoTec, FL, USA; setup B_2). The raw analog video signals of all setups were digitized at a frame rate of 60 Hz on a dedicated PC running a gaze extraction software

(iViewX™, SMI, Germany, and Clearview™, Tobii Technology AB, Sweden, respectively) which produced real-time gaze position output. Careful eyetracking calibration was performed prior to each training or data acquisition session in order to yield valid gaze positions in a stimulus-related coordinate system. Via a fast network connection, gaze position updates were transferred and, thus, made available to another PC running the software which controlled the stimulation paradigm (Presentation™, http://www.neurobs.com).

Interactive eyetracking with overt feedback

This version was established so that study participants could engage in cognitive tasks using their eye movements only while receiving visual feedback. To this end we coupled the volunteer's current gaze position to the location of a cursor-like object on the screen (hereafter: gaze cursor). Using their eyes, the volunteers could, thus, voluntarily move the gaze cursor according to the demands of the tasks. For the automatic detection of gaze fixations on screen targets in real-time the following computer-based algorithm was devised: Using Presentation™ software, gaze positions were transformed into stimulus screen coordinates (pixels). A continuously proceeding 'sliding window' average of the preceding 60 gaze positions was calculated throughout the whole stimulus presentation (Figure 1). In effect, the gaze cursor marked the volunteers' average gaze position within the preceding 1 s time window providing the observer with a smooth gaze-contingent visual stimulus to which the volunteers quickly adapted despite a brief temporal lag. In particular, this procedure lessened blinking artifacts, averaged out fixational eye movements (2–120 arcmin; Martinez-Conde *et al.*, 2004), and attenuated the impact of erroneous gaze estimates caused by intermittent residual imaging artifacts in the eye video signal. Each 'sliding window' average was tested for being part of a coherent fixation period or not and was accepted by the algorithm as part of an ongoing fixation, if the standard deviation of the sliding window gaze elements was below a pre-specified threshold, in which case a counter was incremented. If the standard deviation criterion was not fulfilled, the counter was reset to zero. This procedure was repeated until a fixation period of a pre-specified length, i.e. a pre-specified number of consecutive sliding window averages, was detected. This procedure reliably recognized effective fixations from gaze behavior without prior knowledge of fixation coordinates. Fixations were subsequently tested for being within one of a set of predefined region-of-interests (ROIs) on the stimulus screen. If this was not the case, the algorithm searched for another fixation. This cycle was repeated until either a fixation was found that was within one of the predefined ROIs or the maximum duration of the current task was reached. Time stamps as well as coordinates of detected fixations were stored in a text file for offline data analysis.

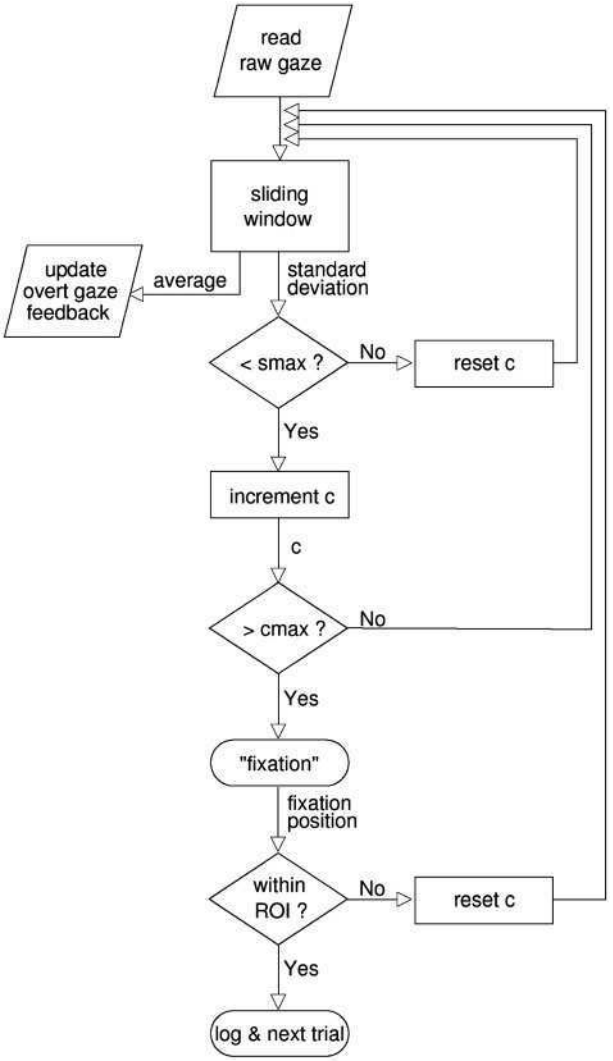

Fig. 1 Flow diagram of gaze data processing. Raw gaze data is stored in a sliding window vector whose average value is shown as a gaze cursor in overt gaze feedback paradigms. A counter (*c*) increments for each consecutive updated sliding window data vector whose standard deviation is below a prespecified threshold (*smax*). The gaze status is identified as 'fixation' if a prespecified number (*cmax*) of consecutive sliding windows is reached. The average gaze position is thereafter tested for being within one of a given set of ROIs. The described procedure is typically run until a fixation was found within one of the given ROIs triggering a step forward in the experimental paradigm, e.g. the presentation of a new visual stimulus or the change of the gaze direction of a virtual character.

Interactive eyetracking with covert feedback

As during interactive eyetracking with overt feedback, here, participants engage in and 'drive' an experimental paradigm by looking at different locations on the screen. In this version of the setup, however, they do not receive visual feedback in form of a gaze cursor.

In conjunction with a virtual character whose gaze behavior could be made contingent upon fixations detected in ROIs this was done to generate an ecologically valid setting in which the gaze behavior of the virtual other could change in response to the human observer's gaze position on the

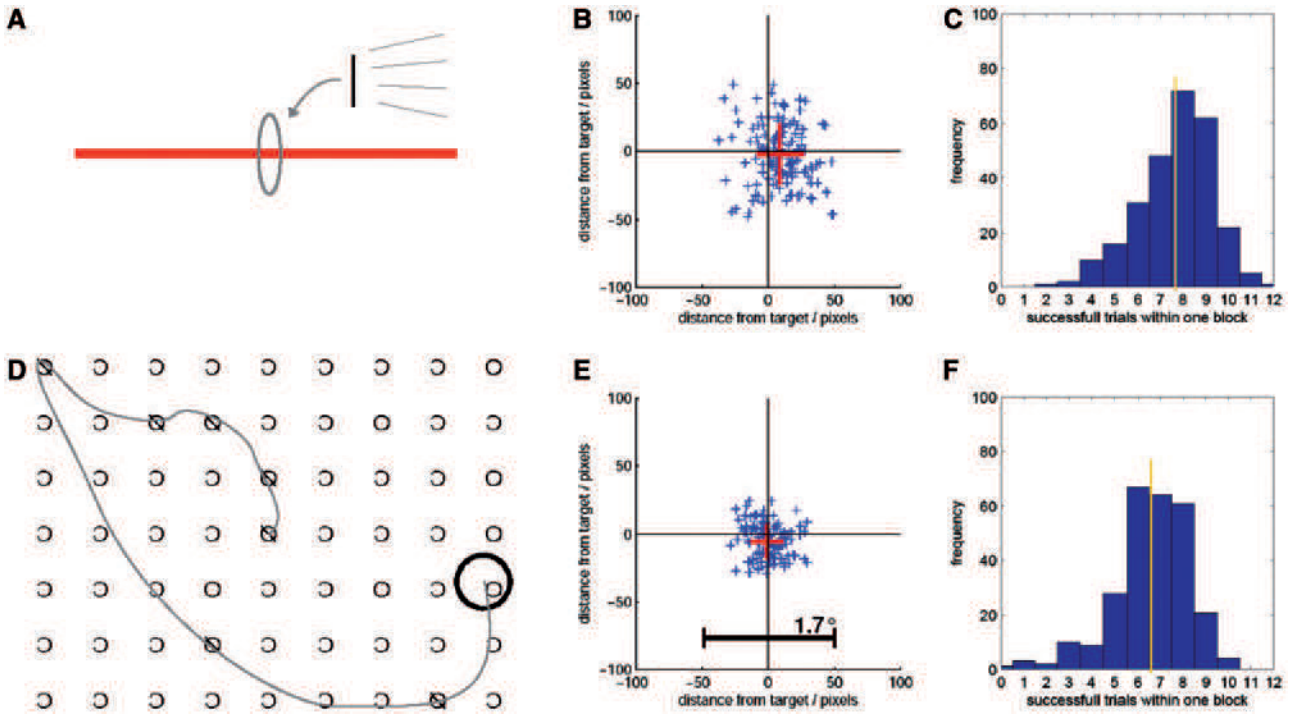

Fig. 2 Overt gaze feedback tasks. (**A**) While being scanned using 'dummy' fMRI scans, a subject was instructed to bisect a horizontal line with a vertical line which is locked to her gaze. In effect, the subject performed the task by fixating the perceived line center. The screen positions of the horizontal lines were randomized. (**B**) Example of the spatial precision of one subject performing the bisection task. Each of the blue symbols represents one bisection position relative to the true line center. The red cross denotes the average and standard deviation of the spatial bisection error. (**C**) This histogram shows the frequency distribution of the number of line bisections that subjects were able to perform within one task/block length (21.9 s). On average over 18 subjects, each performing the tasks 15 times, 7.7 line bisections were successfully performed within one block length revealing a rather fluent task performance. (**D**) In a second task, the subject was asked to cancel targets ('0') among distractors by centering the black circle over each target until marked as cancelled. Since the circle's position was locked to the subject's gaze she only had to find and fixate the targets one by one. (**E**) The spatial precision of the same subject as in (B) performing the cancellation task. (**F**) On average over 18 subjects, 6.6 cancellations were successfully performed within one block.

stimulus screen. In this setup the temporal delay between a relevant fixation and the reaction of the virtual character needed to be small to successfully induce a fluent experience of reciprocal exchange between the participant and the virtual character. Not providing continuous visual feedback via a gaze cursor precluded the possibility of participants being distracted by the gaze cursor but also to adapt to possible measurement errors, even if minimal in size. For example, participants' head movements or variations in eye illumination could invalidate the initial eye tracking calibration and lead to displaced gaze coordinates. This type of technical problem could be met by defining larger ROIs that were less sensitive to distortions in gaze estimations. Despite minimal offset errors in the gaze coordinates, thus, meaningful reactions of the virtual character were still possible.

The tasks inside the MR environment

To make use of the *overt* interactive eyetracking mode we implemented two tasks resembling clinical bedside tests for visuo-spatial neglect. In a line bisection task participants had to bisect a horizontal line by fixating it centrally thereby moving a gaze cursor (a vertical line) on the screen into the desired position (Figure 2). In a target cancellation task volunteers had to search for and single out randomly distributed targets among distractors by fixating them one by one thereby moving a gaze cursor (a circle) over each of the detected targets (see online supplementary data for video of task performance). Participants were informed about the paradigm prior to entering the scanner room. Both overt interactive eyetracking tasks were tested with visual stimulation delivered via the TFT screen (setup A) and via goggles (setups B_1 and B_2). After careful calibration of the eyetracking, volunteers were allowed to get adjusted to the procedure and then went on to perform the task while lying inside the scanner. During this period we ran 'dummy' EPI sequences with MRI parameters identical to those in standard imaging experiments. We thereby introduced EPI artifacts in the eyetracking data to test that the devised algorithm would be able to successfully cope with the added noise.

In order to test the *covert* interactive eyetracking mode we made use of a task in which test subjects were asked to respond to or probe the gaze behavior shown by an anthropomorphic virtual character on screen (Figure 3; Schilbach *et al.*, in press). Before participation test subjects were instructed that the gaze behavior shown by the virtual character on screen was actually controlled by a real person who was also participating in the experiment outside the scanner. Likewise, their own gaze behavior was said to be visualized

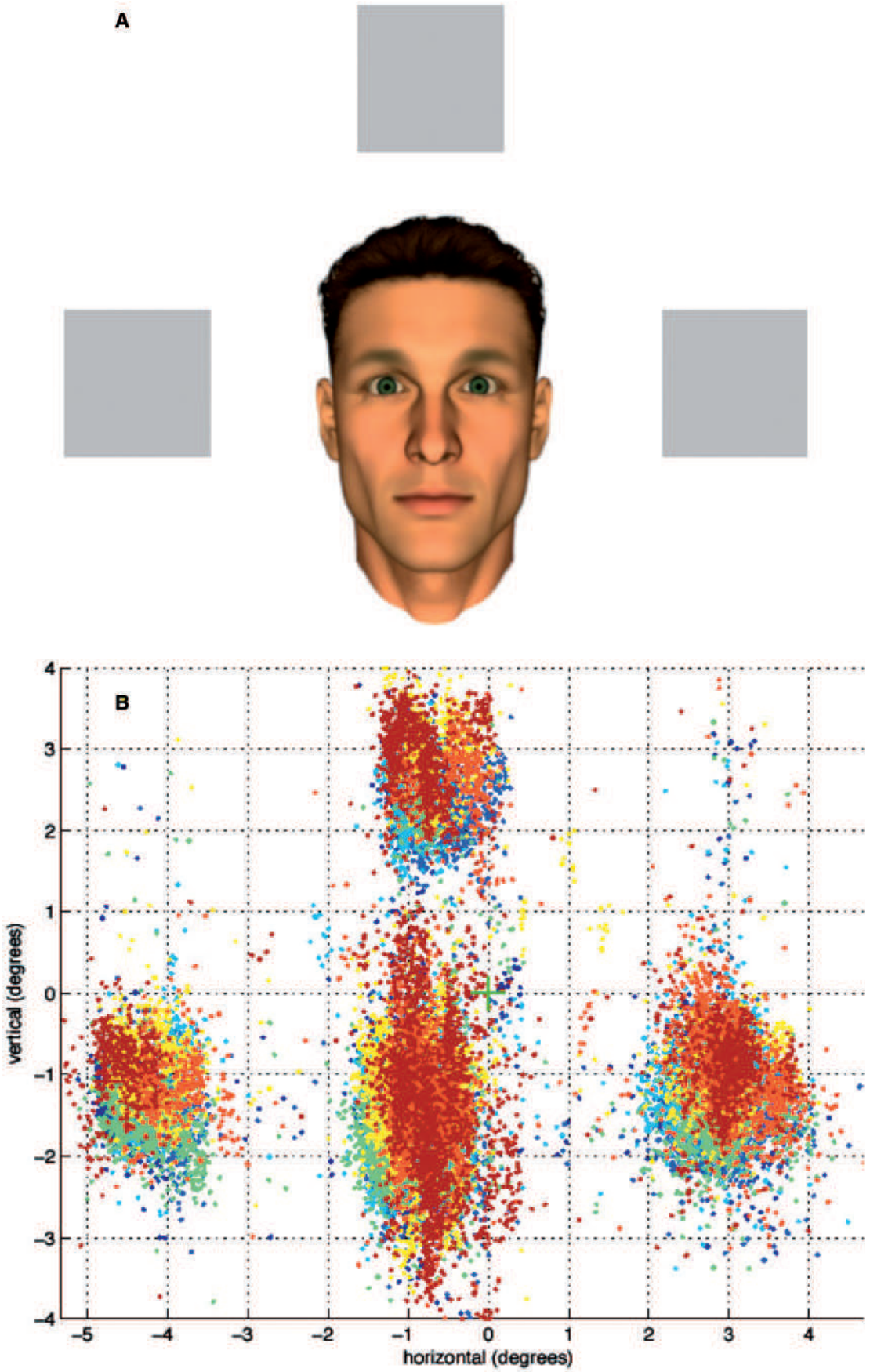

Fig. 3 Covert gaze feedback task. (**A**) Screenshot (as seen by participants) depicting anthropomorphic virtual character and three objects. (**B**) Illustration of gaze samples obtained for one exemplary participant during the experiment.

for the other participant outside the MR scanner, so that both participants could engage in gaze-based 'online' interaction. During functional neuroimaging participants were instructed to direct the gaze of the other person towards one of three objects by looking at it. Alternatively, they were asked to respond to gaze shifts of the virtual character by either following or not following them to then look at an object. The gaze behavior of the other was made contingent to the participant's gaze and systematically varied in a 2×2 factorial design (joint attention *vs* non-joint attention; self-initiation *vs* other-initiation; see online supplementary data for video of task performance). The neural correlates of task performance were investigated employing fMRI in 21 participants.

In this task, stimuli were presented to the participants lying inside the MR scanner using setup A. Due to the screen's distance from the volunteers' eyes and the corresponding narrow field of view; changes of the virtual character's gaze behavior were easily observable while focusing on one of the three objects. Functional MRI (fMRI) data was acquired on a Siemens Trio 3T whole-body scanner (Erlangen, Germany) using blood-oxygen-level-dependent (BOLD) contrast (Gradient-echo EPI pulse sequence, TR = 2.304 s, slice thickness 3 mm, 38 axial slices, in-plane resolution 3×3 mm). Additional high-resolution anatomical images (voxel size $1 \times 1 \times 1$ mm^3) were acquired using a standard T_1-weighted 3D MP-RAGE sequence. The neuroimaging data was preprocessed and analyzed using a general linear model (GLM) as implemented in SPM5 (for further details see: Schilbach *et al.*, in press).

RESULTS

Interactive eyetracking under the constraints of fMRI was successfully installed in all setups. The quality depended on the length of the sliding window which in turn depended on the raw gaze data variance, the main source of which were residual imaging artifacts. Fixational eye movements were the other important source of raw gaze data jitter. The amplitude of such eye movements are in the range of 2–120 arcmin (Martinez-Conde *et al.*, 2004) and their size in screen pixel coordinates scales with the eye's distance to the screen display. This rendered the fixational eye movements' impact particularly disadvantageous in setup A where a screen distance of 245 cm translated the range of 2–120 arcmin to 2–120 pixels on the screen (given a 800×600 pixels resolution). On the other hand, setup A did also include important advantages regarding the handling of the eyetracking camera compared to setups using goggles. The distance as well as the angle with which the camera was positioned in relation to the test subject's eyes region could be adjusted more easily when the eyetracking camera was mounted to the head coil as in the TFT-based setup. However, goggle systems that allowed a precise eyetracking with their built-in camera did hardly need any camera adjustments (e.g. setup B_2). After the feasibility of all setups was secured we chose to use our 3 T MRI system and opted for the TFT-based visual stimulation setup for data acquisition. We nevertheless stress, that the other setups allowed running the paradigms as well.

Interactive eyetracking with overt feedback (line bisection and target cancellation)

During the task in which overt feedback was given, participants were able to use the gaze cursor which they could move across the screen in concordance with their eye movements to execute the task. For the bisection task this amounted to subjects completing an average of 7.7 line bisections in blocks of 21.9 s ($n = 18$ subjects), whereas during the cancellation task subjects were able to cancel an average of 6.6 targets within the same time (Figure 2). Subjects needed an average of 2.8 s to judge the center of a given line and position the gaze cursor in the respective position. An average additional 0.5 s was needed to search, find, and position the gaze cursor in cancellation tasks (average time between cancellations: 3.3 s). The time subjects needed to choose targets generally depended on the length of the sliding window. If the sliding window length was too small, increased residual gaze cursor jittering made it hard for volunteers to 'focus' a target, whereas too long a sliding window increased the temporal lag of the gaze cursor, which reduced the intuitive usability of the gaze feedback. Apart from this, the spatial precision of target choices was increased when the eye tracking calibration was optimal and subjects were well adjusted to the temporal lag of the gaze cursor (Figure 2).

Interactive eyetracking with covert feedback (joint attention)

Having been able to use overt gaze feedback successfully to drive an experiment, we went on to perform the joint attention task which included running regular fMRI measurements. During this task subjects were not given visual feedback in the form of a gaze cursor because we wanted to create a naturalistic as possible setup which allowed for an immersive experience during which participants could interact with the virtual other similarly to how one might interact with another person by means of gaze behavior in real life. In spite of the absence of continuous visual feedback participants were able to fulfill the task in which they had been asked to engage: they were able to establish 'eye-contact' with the virtual character and to respond to the virtual character either by following or not following its gaze to either fixate one of three visible objects 'together with' the virtual character (other-initiated joint attention: OTHER_JA) or not (other-initiated nonjoint attention: OTHER_NOJA; Figure 3A). Conversely, they were also able to establish 'eye-contact' and subsequently direct the virtual character's gaze towards one of the three objects (self-initiated joint attention: SELF_JA). In an equal number of occasions subjects were unable to do so as the character would 'react' by fixating an object other than the one chosen by the

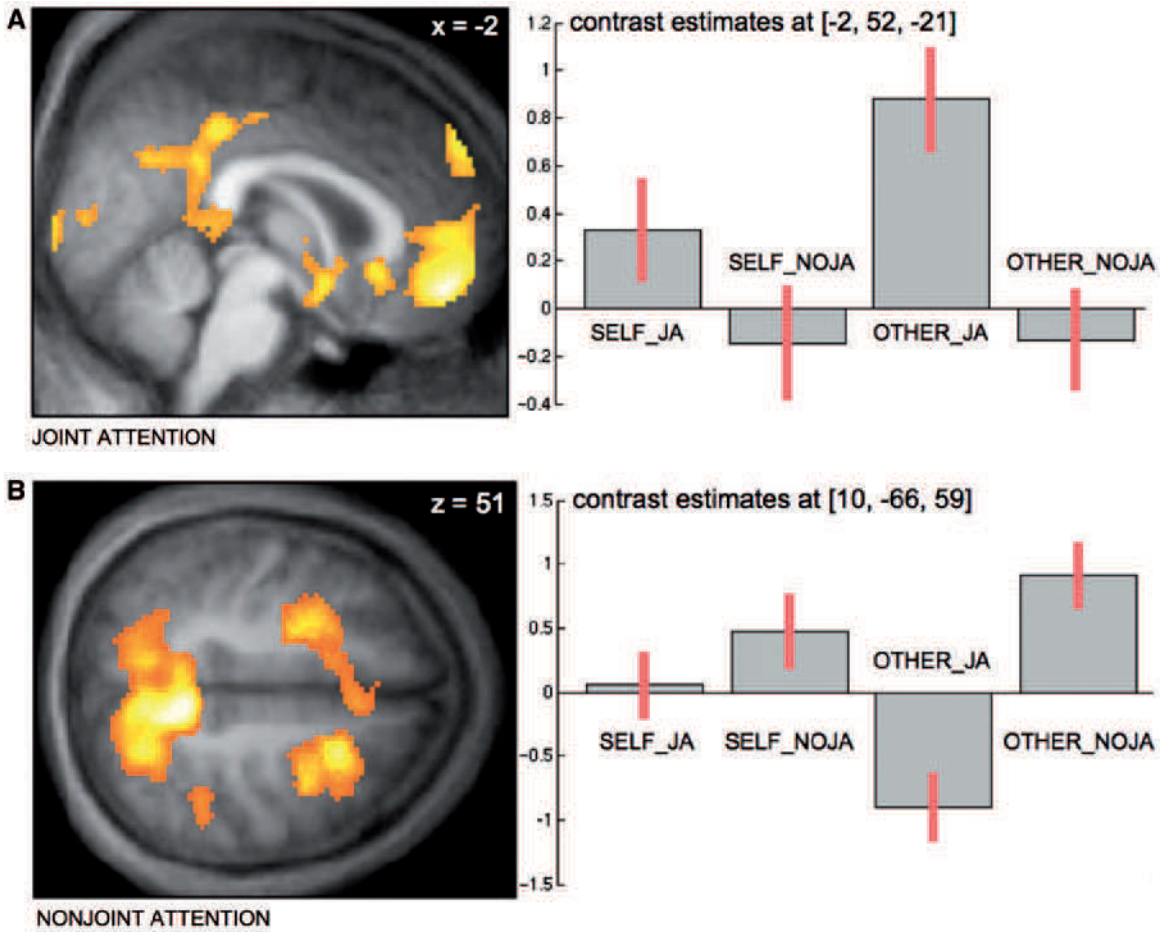

Fig. 4 Neural correlates of joint attention task. (**A**) Differential increase of neural activity in MPFC, PCC as well as ventral striatum and anterior temporal poles (latter not illustrated here) for main effect of joint attention. (**B**) Differential increase of neural activity in medial and lateral parietal as well as frontal cortex bilaterally for main effect of nonjoint attention (taken from: Schilbach *et al.*, in press).

participant (self-initiated nonjoint attention: SELF-NOJA). On average, this procedure amounted to approximately four object fixations per 'interaction segment' (i.e. a block of 18 s duration) for all conditions (SELF_JA: 4.10 ($n = 21$, s.d. = 0.68), SELF_NOJA: 4.06 ($n = 21$, s.d. = 0.79), OTHER_JA: 3.96 ($n = 21$, s.d. = 0.88), OTHER_NOJA: 4.03 ($n = 21$, s.d. = 0.88); see Figure 3B for exemplary gaze data; for more details see Schilbach *et al.*, in press). FMRI results demonstrated, firstly, that interpersonal gaze coordination and 'joint attention' (main effect of JA) resulted in a differential increase of neural activity in the medial prefrontal cortex (MPFC) and posterior cingulate cortex (PCC) as well as the anterior temporal poles bilaterally. While this activation pattern bears some resemblance with the 'default mode of brain function' (Raichle *et al.*, 2001; Schilbach *et al.*, 2008), activations in ventral and dorsal MPFC—at voxel-level correction for multiple comparisons—have been related to outcome monitoring and the understanding of communicative intent (Amodio and Frith, 2006) as well as representations of triadic relations (Saxe, 2006). Conversely, looking at an object different from the one attended by the virtual character—regardless of whether or not this was self-initiated (main effect of NOJA)—recruited a bilateral fronto-parietal network known to be involved in attention and eye-movement control (Schilbach *et al.*, in press; Figure 4).

DISCUSSION

Here, we have presented a method by which eyetracking data obtained from study participants lying inside an MR scanner can be processed in real-time in order to directly influence visual stimulus material in spite of the electromagnetic noise associated with fMRI measurements. This can be realized by using a form of *overt* gaze feedback, i.e. a gaze cursor which subjects can control with their eye movements to carry out a task. Alternatively, this can be done by means of *covert* gaze feedback where gaze data is used to systematically manipulate and drive the visual stimulation unbeknownst to the participant. In combination with the presentation of anthropomorphic virtual characters whose behavior can be made responsive to the participants' fixations, the latter technique can be used to allow participants to engage in reciprocal 'online' interaction with a virtual other similar to instances of interpersonal gaze coordination in real-life social

encounters. Even though similar approaches of using gaze-contingent stimuli have been used in other areas of research (e.g. Duchowski *et al.*, 2004), the development of an MR-compatible version of the technique as presented here is of crucial importance for social cognitive and affective neuroscience as it allows to experimentally target the neural underpinnings of processes during 'online' interaction which have so far been largely inaccessible due to the technical constraints of the MR environment.

Different possible paradigms come to mind which could benefit from making use of the here described method. Given the scope of this article, we will limit our description to interactive paradigms in which gaze behavior is exchanged between a human observer and a virtual character. We will focus here on *dyadic interaction* between two interactors ('Me–You'), but also on *triadic interaction* where two interactors relate to an object in the environment ('Me–You–This'; Saxe, 2006).

Within dyadic interaction gaze is known to have important regulatory functions impacting on a wide range of cognitive, affective and motivational processes (Argyle and Cook, 1976; Emery, 2000). Furthermore, gaze is known to influence our social perception and evaluation of others (e.g. Macrae *et al.*, 2002; Mason *et al.*, 2005) as it conveys the direction of an agent's attention and has been suggested to be closely related to mentalizing, i.e. our ability to understand other people's mental states (Nummenmaa and Calder, 2009). Importantly, gaze is also known to 'connect' human beings in everyday life situations by means of a 'communication loop' in which interactors impact reciprocally on each others' behavior (e.g. Frith, 2007, p. 175). This procedural dimension of 'social gaze' in 'online' interaction has only recently begun to be systematically investigated (e.g. Senju and Csibra, 2008) and promises to allow radically new insights into the temporal dynamics of implicit interpersonal 'alignment'.

As many previous studies concerning the social effects of gaze on person perception have used static, non-interactive stimuli, it may be important to revisit these paradigms by making use of this new technique to validate whether the findings actually result from social communicative processes or not. Here, it is important to note, however, that our setup in its current version does not allow to investigate real social interaction (as in the setup used by Montague *et al.*, 2002), but uses anthropomorphic virtual characters in conjunction with a cover story to generate the impression of interacting with a 'mindful' agent. While this can be seen as a limitation of our setup, it is important to note that using gaze feedback has the important benefit of enabling the systematic study of a major component of real interpersonal interaction as it provides a naturalistic way to engage participants. Furthermore, future research could explicitly address how variations of the temporal and stochastic characteristics of a virtual character's gaze behavior made contingent upon the human observer's gaze impact on a human observer's perception of the nature of the agent ('social' Turing test).

Apart from aspects related to dyadic interaction, gaze is also known to contribute to the establishment of triadic relations between two interactors who can look at an object together and engage in (gaze-based) joint attention (Moore and Dunham, 1995). Apart from the convergence of gaze directions, this, importantly, also requires mutual awareness of being intentionally directed towards the same aspect of the world which may result directly from the process of interaction. Therefore, joint attention can be construed as an interactively constituted phenomenon whose different facets can only be explored by making use of an interactive paradigm (e.g. Schilbach *et al.*, in press). Interestingly, it has been suggested that being actively engaged in triadic interaction may have an impact both on the perception of the other person (e.g. his/her trustworthiness and attractiveness) as well as on the perception of an object (e.g. its value) that may be jointly attended (Heider, 1958).

To the best of our knowledge, there are no neuroimaging studies which have targeted the neural correlates of the perception of jointly attended objects. Such investigations might be extremely informative, however, by allowing the empirical investigation of the neural correlates of different formats or varying degrees of shared intentionality and could help to investigate the complex relationship of implicit and explicit processes involved in social cognition (Frith and Frith, 2008). Further applications of the method could include investigations of how interactive gaze cues shown by a virtual character impact on object-related decision-making or memory performance. Finally, gaze-based triadic interaction could also help to disentangle differences between 'online' and 'offline' social cognition, i.e. social cognition from a second- or third-person-perspective, by realizing interactions between two virtual agents and a human observer while introducing systematic differences in social responsiveness of the agents making them more or less likely to actually engage with the participant (Eisenberger *et al.*, 2003).

Apart from investigations of dyadic and triadic interaction in healthy adults, we also see great potential in using the above described method to investigate social cognition in development and diagnostic groups. In spite of the importance of joint attention in ontogeny the neural correlates of this significant phenomenon are only incompletely understood. Given that our paradigm does not rely upon verbal information and higher-order reasoning about others' mental states, but relies upon naturally occurring social behavior, it might prove to be particularly helpful for the study of the neurofunctional substrates of the development of social cognitive and perceptual abilities during ontogeny.

Specific alterations of social cognition are known to be characteristic of psychiatric disorders such as autism and schizophrenia. In the former case, a dissociation between implicit and explicit processes underlying social cognition

has recently been emphasized (Senju *et al.*, 2009). Also, it has been suggested that autistic individuals might be more sensitive to perfect, non-social as compared to imperfect, social contingencies in the environment (Gergely, 2001; Klin *et al.*, 2009). We suggest that the investigation of the neural mechanisms underlying these clinically relevant differences in high-functioning autism will benefit substantially from the method described here (see also Boraston and Blakemore, 2007).

CONCLUSIONS

Taken together, we have shown that the interactive eyetracking method here presented allows to generate gaze-contingent stimuli during fMRI in spite of the electromagnetic noise introduced by such measurements. Used in conjunction with anthropomorphic virtual characters whose behavior can be made 'responsive' to the participant's current gaze position, this method has the potential to substantially increase our knowledge of the neural mechanisms underlying social cognition by making psychological processes accessible for empirical investigation that are involved in the interpersonal coordination of gaze behavior, both in dyadic and triadic interaction. Making use of this new 'tool of the trade', we suggest, could open up an entire new avenue of research in social cognitive and affective neuroscience.

SUPPLEMENTARY DATA

Supplementary data are available at *SCAN* online.

REFERENCES

Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Review Neuroscience*, *7*(4), 268–77.

Argyle, M., Cook, M. (1976). *Gaze and Mutual gaze*. Cambridge: Cambridge University Press.

Bailenson, J.N., Blascovich, J., Beall, A.C., Loomis, J.M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin*, *29*(7), 819–33.

Becchio, C., Bertone, C., Castiello, U. (2008). How the gaze of others influences object processing. *Trends in Cognitive Science*, *12*(7), 254–8.

Bente, G., Krämer, N.C., Petersen, A., de Ruiter, J.P. (2001). Computer animated movement and person perception: methodological advances in nonverbal behavior research. *Journal of Nonverbal Behavior*, *25*(3), 151–66.

Bente, G., Krämer, N.C. (2002). Virtual gestures: Analyzing social presence effects of computer-mediated and computer-generated nonverbal behaviour. In: Gouveia , Feliz, R., editors. *Proceedings of the Fifth Annual International Workshop PRESENCE, October, 2002, Porto, Portugal.* http://www.temple.edu/ispr/prev_conferences/proceedings/2002/Final%20papers/Bente%20and%20Kramer.pdf (15 Dec 2009, date last accessed).

Bente, G., Eschenburg, F., Aelker, L. (2007). Effects of simulated gaze on social presence, person perception and personality attribution in avatar-mediated communication. *Proceedings of the 10th Annual International Workshop on Presence 2007.* http://www.temple.edu/ispr/prev_conferences/proceedings/2007/Bente,%20Eschenburg,%20and%20Aelker.pdf (15 Dec 2009, date last accessed).

Bente, G., Krämer, N.C., Eschenburg, F. (2008a). Is there anybody out there? Analyzing the effects of embodiment and nonverbal behavior in avatar-mediated communication. In: Konijn, E., Utz, S., Tanis, M., Barnes, S., editors. *Mediated Interpersonal Communication*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 131–57.

Bente, G., Senokozlieva, M., Pennig, S., Al Issa, A., Fischer, O. (2008b). Deciphering the secret code: A new methodology for the cross-cultural analysis of nonverbal behavior. *Behavior Research Methods*, *40*(1), 269–77.

Boraston, Z., Blakemore, S.J. (2007). The application of eye-tracking technology in the study of autism. *Journal of Physiology*, *581*(Pt 3), 893–8.

Duchowski, A.T., Cournia, N., Murphy, H. (2004). Gaze-contingent displays: a review. *Cyberpsychol Behav*, *7*(6), 621–34.

Eisenberger, N.I., Lieberman, M.D., Williams, K.D. (2003). Does rejection hurt? An FMRI study of social exclusion. *Science*, *302*(5643), 290–2.

Emery, N.J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience Biobehavior Review*, *24*(6), 581–604.

Frith, C.D. (2007). *Making up the mind*. Blackwell Publishing.

Frith, C.D., Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron*, *60*(3), 503–10.

Frith, U., Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, *358*(1431), 459–73.

Gergely, G. (2001). The obscure object of desire: 'nearly, but clearly not, like me': contingency preference in normal children versus children with autism. *Bull Menninger Clin*, *65*(3), 411–26.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Keysers, C., Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends in Cognitive Science*, *11*(5), 194–6.

Klin, A., Jones, W., Schultz, R., Volkmar, F. (2003). The enactive mind, or from actions to cognition: lessons from autism. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, *358*(1430), 345–60.

Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*, *459*(7244), 257–61.

Kuzmanovic, B., Georgescu, A.L., Eickhoff, S.B., et al. (2009). Duration matters: dissociating neural correlates of detection and evaluation of social gaze. *Neuroimage*, *46*(4), 1154–63.

Lieberman, M.D. (2007). Social cognitive neuroscience: a review of core processes. *Annual Review of Psychology*, *58*, 259–89.

Macrae, C.N., Hood, B.M., Milne, A.B., Rowe, A.C., Mason, M.F. (2002). Are you looking at me? Eye gaze and person perception. *Psycholohical Science*, *13*(5), 460–4.

Martinez-Conde, S., Macknik, S. L., Hubel, D.H. (2004). The role of fixational eye movements in visual perception. *Nature Review of Neuroscience*, *5*(3), 229–40.

Mason, M.F., Tatkow, E.P., Macrae, C.N. (2005). The look of love: gaze shifts and person perception. *Psychological Science*, *16*(3), 236–9.

Montague, P.R., Berns, G.S., Cohen, J.D., et al. (2002). Hyperscanning: simultaneous fMRI during linked social interactions. *Neuroimage*, *16*(4), 1159–64.

Moore, C., Dunham, P.J. (1995). *Joint Attention: Its Origin and Role in Development*. Hillsdale: Lawrence Erlbaum.

Nummenmaa, L., Calder, A.J. (2009). Neural mechanisms of social attention. *Trends in Cognitive Science*, *13*(3), 135–43.

Park, K.M., Kim, J.J., Ku, J., et al. (2009). Neural basis of attributional style in schizophrenia. *Neuroscience Letters*, *459*(1), 35–40.

Pelphrey, K.A., Morris, J.P., McCarthy, G. (2005). Neural basis of eye gaze processing deficits in autism. *Brain*, *128*, 1038–48.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L. (2001). A default mode of brain function. *Proceedings of National Academy of Sciences United States of America*, *98*(2), 676–82.

Reddy, V. (2003). On being the object of attention: implications for self-other consciousness. *Trends in Cognitive Sciences*, *7*(9), 397–402.

Sanchez-Vives, M.V., Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Review of Neurosciences*, *6*(4), 332–9.

Saxe, R. (2006). Uniquely human cognition. *Current Opinion of Neurobiology*, *16*, 235–9.

Sanfey, A.G. (2007). Social decision-making: insights from game theory and neuroscience. *Science*, *318*(5850), 598–602.

Schilbach, L., Wohlschlaeger, A.M., Kraemer, N.C., et al. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, *44*(5), 718–30.

Schilbach, L., Eickhoff, S.B., Rotarska-Jagiela, A., Fink, G.R., Vogeley, K. (2008). Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the ''default system'' of the brain. *Conscious Cognition*, *17*(2), 457–67.

Schilbach, L., Wilms, M., Eickhoff, S.B., et al. (in press). Minds made for sharing. Initiating joint attention recruits reward-related neurocircuitry. *Journal of Cognitive Neuroscience*, [Epub-ahead of print; 25 November 2009]; doi: 10.1162/jocn.2009.21401.

Senju, A., Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*(9), 668–71.

Senju, A., Southgate, V., White, S., Frith, U. (2009). Mindblind eyes: an absence of spontaneous theory of mind in Asperger syndrome. *Science*, *325*(5942), 883–5.

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neurosci Biobehaviour Review*, *30*(6), 855–63.

Slater, M., Pertaub, D.P., Barker, C., Clark, D.M. (2006). An experimental study on fear of public speaking using a virtual environment. *Cyberpsychol Behaviour*, *9*(5), 627–33.

Spiers, H.J., Maguire, E.A. (2006). Spontaneous mentalizing during an interactive real world task: an fMRI study. *Neuropsychologia*, *44*(10), 1674–82.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 03 December 2012
doi: 10.3389/fpsyg.2012.00537

Eyes on the mind: investigating the influence of gaze dynamics on the perception of others in real-time social interaction

Ulrich J. Pfeiffer[1,2]*, Leonhard Schilbach[1,3], Mathis Jording[1], Bert Timmermans[1], Gary Bente[4] and Kai Vogeley[1,2]

[1] *Neuroimaging Group, Department of Psychiatry, University Hospital Cologne, Cologne, Germany*
[2] *Institute of Neuroscience and Medicine – Cognitive Neurology (INM3), Research Center Juelich, Juelich, Germany*
[3] *Max-Planck-Institute for Neurological Research, Cologne, Germany*
[4] *Faculty of Human Sciences, Department of Media and Social Psychology, University of Cologne, Cologne, Germany*

Edited by:
Judith Holler, Max Planck Institute Psycholinguistics, Netherlands

Reviewed by:
Daniel Richardson, University College London, UK
Helene Kreysa, Friedrich-Schiller-University of Jena, Germany

****Correspondence:***
Ulrich J. Pfeiffer, Neuroimaging Group, Department of Psychiatry, University Hospital Cologne, Kerpener Strasse 62, 50937 Cologne, Germany.
e-mail: ulrich.pfeiffer@uk-koeln.de

Social gaze provides a window into the interests and intentions of others and allows us to actively point out our own. It enables us to engage in *triadic* interactions involving human actors and physical objects and to build an indispensable basis for coordinated action and collaborative efforts. The object-related aspect of gaze in combination with the fact that any motor act of looking encompasses both input and output of the minds involved makes this non-verbal cue system particularly interesting for research in embodied social cognition. Social gaze comprises several core components, such as gaze-following or gaze aversion. Gaze-following can result in situations of either "joint attention" or "shared attention." The former describes situations in which the gaze-follower is aware of sharing a joint visual focus with the gazer. The latter refers to a situation in which gazer and gaze-follower focus on the same object *and* both are aware of their reciprocal awareness of this joint focus. Here, a novel interactive eye-tracking paradigm suited for studying triadic interactions was used to explore two aspects of social gaze. Experiments 1a and 1b assessed how the latency of another person's gaze reactions (i.e., gaze-following or gaze version) affected participants' sense of agency, which was measured by their experience of relatedness of these reactions. Results demonstrate that both timing and congruency of a gaze reaction as well as the other's action options influence the sense of agency. Experiment 2 explored differences in gaze dynamics when participants were asked to establish either joint or shared attention. Findings indicate that establishing shared attention takes longer and requires a larger number of gaze shifts as compared to joint attention, which more closely seems to resemble simple visual detection. Taken together, novel insights into the sense of agency and the awareness of others in gaze-based interaction are provided.

Keywords: gaze-following, joint attention, shared attention, social interaction, agency, mentalizing, eye-tracking

INTRODUCTION

The visual system is a major source of information about the environment. In face-to-face social encounters it is not only a source of information but also a crucial means of non-verbal communication. Imagine the following everyday situation: you are sitting at the bar of a pub gazing contemplatively at your empty glass. Suddenly the bartender walks by and observes that your eyes are directed at the empty glass. As soon as you direct your gaze at him and back to the glass he will – without words – understand that you need another drink. Such instances of "social gaze" demonstrate how meaning can be conveyed by simple acts of looking. A considerable amount of research has been devoted to the development and function of social gaze (Argyle and Cook, 1976; Mundy and Newell, 2007; Shepherd, 2010). Gaze represents a non-verbal cue system which reflects perception and action simultaneously, or in which, as Gibson and Pick, 1963, p. 368) have noted, "any act of looking can be treated as a source of stimulation as well as a type of response." Its salience in social encounters makes gaze a perfect tool to study "online" social interaction, i.e., face-to-face interaction between two persons in real-time (Schilbach et al., 2011).

Mainly due to methodological constraints, the study of online interaction has largely been neglected by researchers in social cognition (Schilbach et al., in press). In recent years, however, there have been exciting advances to create tools for the investigation of non-verbal and especially gaze-based social interaction (Redcay et al., 2010; Wilms et al., 2010; Staudte and Crocker, 2011; Bayliss et al., 2012). For example, Redcay et al. (2010) established a setup in which participants inside an MRI scanner could either interact face-to-face with an experimenter via a live video feed or watch a recording of the experimenter's behavior during previous interactions, thereby enabling the investigation of the processing of dynamic features of social interaction. Staudte and Crocker (2011) designed a series of experiments in which

participants interacted with an artificial agent (i.e., a robot) in order to study the dynamic coupling between gaze and language in verbal human-robot interaction. Recently, Wilms et al. (2010) introduced an interactive eye-tracking setup which allows participants to interact with an anthropomorphic virtual character in a gaze-contingent manner. A similar program has been created recently by another group to study face-to-face interaction in social contexts (Grynszpan et al., 2012).

The advent of virtual reality techniques for research in neuroscience and psychology (Tarr and Warren, 2002; Bohil et al., 2011) has raised the general question why we need these displays to study human cognition. Bohil et al. (2011, p. 752) have noted that "an enduring tension exists between ecological validity and experimental control" in psychological research. They suggest that virtual reality techniques provide a way out of this dilemma because they provide naturalistic, real-world-like displays whilst offering full control over a selected set of experimental variables. Indeed, studies addressing the validity of using virtual characters have demonstrated that the interaction with virtual agents elicits social behaviors which are similar to real interaction (von der Pütten et al., 2010) and that uncontrolled aspects of another person's outer appearance and non-verbal behavior can be filtered out while participants' overall impression of an interaction remains intact (Vogeley and Bente, 2010). In addition, avatar- and video-mediated communication have shown to create comparable levels of experienced social presence and intimateness (Bente et al., 2008).

Before such paradigms can be used to study gaze in more complex social scenarios, basic parameters of different processes of social gaze need to be identified. Several of these processes have been defined by Emery (2000): *direct (or mutual) gaze* – a situation where two individuals direct their gaze at each other – is described as the most basic process of social gaze. If one individual detects that the other averts its gaze this can serve as a cue for a *gaze-following* reaction to the other's novel focus of visual attention. This results in a situation of *joint attention (JA)*, in which the gaze-follower is aware that he and the gazer have the same focus of attention – for instance, an object in the environment. In other words, in JA another person's gaze is hence used as a cue to this person's visual attention. This has been argued to represent a crucial prerequisite for the gaze-follower to infer the gazer's mental states (e.g., thoughts, intentions, feelings. . .) regarding an object of joint focus (Gopnik et al., 1994), an ability commonly referred to as mentalizing (Frith and Frith, 2006). Notably, JA does not require the gazer to be aware of the gaze-follower's reaction. In contrast, *shared attention (SA)* requires that *both* individuals are aware of focusing on the same object *and* of each other's reciprocal awareness of this joint attentional focus (Emery, 2000). Moreover, SA has been argued (Moll and Tomasello, 2007) to involve the gazer's intention to direct the other's gaze to a certain object in order to achieve a shared goal or share an experience, thereby providing a behaviorally accessible measure of shared intentionality. Notably, different but often overlapping descriptions of JA or SA exist in the literature (e.g., Clark, 1996; Povinelli and Eddy, 1996; Tomasello et al., 2005; Frischen et al., 2007; Mundy and Newell, 2007). The study presented in this article is largely guided by the comparably mechanistic account of Emery (2000), which provides a clear conceptual distinction between JA and SA that is suited to provide empirical access to these processes.

Joint and shared attention constitute so-called triadic social interactions. In contrast to dyadic interactions which develop early in infancy and involve processes such as mutual gaze or reciprocal emotional displays (Stern, 1974), triadic interactions are characterized by involving "the referential triangle of child, adult, and some third event or entity to which the participants share attention" (Carpenter et al., 1998, p. 1). The establishment of reference to a certain aspect of the environment in a triadic interaction thus creates a form of perceptual common ground (Clark, 1996). This is a prerequisite for understanding each other's goals and intentions regarding the object of joint focus. So far, however, the temporal and spatial dynamics of gaze in triadic interactions have not been studied systematically using interactive (i.e., gaze-contingent) paradigms (for discussion, see Becchio et al., 2010; Schilbach et al., in press). Although pictures of objects have been used in gaze cueing studies (Bayliss et al., 2006, 2007; van der Weiden et al., 2010), interactive eye-tracking studies so far have been limited to simple geometric shapes as stimuli (Schilbach et al., 2010; Wilms et al., 2010; Pfeiffer et al., 2011).

Using pictures of real-world objects, the current study employs a more ecologically valid interactive eye-tracking setup to address the following questions: (1) *How does the perception of JA depend on the congruency (i.e., gaze-following and gaze aversion) and latency of another person's gaze reactions?* In experiments 1a and 1b, the effect of the congruency of gaze reactions – gaze-following and gaze aversion – as well as the latency with which these reactions follow participants' gaze shifts was manipulated. To this end, participants interacted with a virtual character in brief triadic interactions in which the character would either engage in joint or in non-joint attention (NJA) with different latencies. After each reaction, participants had to indicate how related they experienced this reaction to their own behavior. We argue that this can be taken as a measure to which *degree* participants experienced agency, i.e., that the other's reaction is a consequence of their own action. In its prevalent definition, the sense of agency is described as an all-or-none phenomenon relating to the awareness that we are the initiators of our own actions (de Vignemont and Fourneret, 2004; Synofzik et al., 2008). However, the sense of agency also encompasses an awareness of the consequences (e.g., another person's gaze shifts) inextricably linked to our actions (Bandura, 1989; Pacherie, 2012). As put forward by Pacherie (2012), in social interactions agency experience is not only influenced by high-level cognitive factors and sensorimotor cues, but also by perceptual consequences of one's own actions, including the reactions of another person. Specifically, we hypothesize that participants experience gaze-following (which results in JA) as more strongly related to their own gaze behavior as compared to gaze aversion (which results in disparate attention). It is also predicted that the latency of gaze reactions modulates this experience: very short latencies, which might create an experience of coincidental looking, as well as very long latencies, which might disrupt the temporal contingency between actions, were supposed to decrease participants' sense of agency. (2) *Does gaze behavior differ in situations of JA and SA*? Although the concepts of JA and SA are theoretically distinct, it has never been tested experimentally whether they correspond to

differences in the dynamics of gaze behavior. In Experiment 2, participants engaged in a series of triadic interactions in which they were asked to indicate whenever they experienced JA or SA. We hypothesized that SA requires an increased number of gaze shifts and takes longer to establish as compared to JA.

MATERIALS AND METHODS

In this section, three different experiments will be described. These experiments largely rely on the same materials and methods. For the sake of brevity, those materials and methods that are common to all experiments will be indicated before the procedure of each experiment will be described separately.

PARTICIPANTS

In sum, 95 healthy female and male persons aged 19–42 years ($M = 25.86$, SD = 6.23), with no record of neurologic or psychiatric illnesses volunteered for the study. The numbers for each individual experiment are given in the description of that particular experiment below. All participants were naïve to the scientific purpose of the study and were compensated for their participation (10 Euro/h). Prior to the experiment, participants were asked to sign a written consent form in which they approved that participation is voluntary and that data are used in an anonymized fashion for statistical analysis and scientific publication. The study followed the WMA Declaration of Helsinki (Ethical Principles for Medical Research Involving Human Subjects) and was presented to and approved by the ethics committee of the Medical Faculty of the University Hospital Cologne, Germany.

SETUP AND MATERIALS

We made use of an interactive eye-tracking program recently developed (Wilms et al., 2010). This method allows participants to interact with an anthropomorphic virtual character by means of their eye-movements. Using a high resolution eye-tracking device (Tobii™T1750 Eye-Tracker, Tobii Technology AB, Sweden) with a digitization rate of 50 Hz and an accuracy of 0.5°, participants' eye-movements could be detected exactly. Stimuli were presented on the 17″ TFT screen of the eye-tracker with screen resolution set to 1024 by 768 pixels. Both the participant and the confederate were seated at a distance of 80 cm from their respective eye-tracker as depicted in **Figure 1A**. The viewing angle subtended 32° × 24°. A PC with a dual-core processor and a GeForce 2 MX graphics board controlled the eye-tracker as well as stimulus presentation at a frame rate of 100 Hz. Integrated gaze extraction software (Clearview™, Tobii Technology AB, Sweden) made data available for real-time computation of stimulus presentation to the software package Presentation (Presentation™)[1] which was used to control stimulus presentation in a gaze-contingent manner (for details on the algorithm see Wilms et al., 2010). All data were analyzed using PASW Statistics 20 (SPSS Inc., Chicago, IL, USA)[2].

STIMULI

One male and one female anthropomorphic virtual character were used in this study (Schilbach et al., 2010; Pfeiffer et al., 2011). Except for their eyes, the facial features of these characters were static in order to prevent the influence of non-verbal information other than gaze. Male participants interacted with the male character (exemplarily depicted in **Figure 1B**) and female participants with the female character, respectively. The potency of virtual characters to elicit social presence and the advantages of their usage in experiments on social cognition has been demonstrated previously (for detailed discussion, see Loomis et al., 1999; Bailenson et al., 2003; Vogeley and Bente, 2010).

The 32 object stimuli used here were taken from a previously published study (Bayliss et al., 2006) and consist of two different categories of everyday-life objects, i.e., typical "kitchen" and "garage" objects (**Figure 1B**). They were standardized with respect to likeability ($M = 4.75$, SD = 0.97 on a nine-level scale) and to participants' ability to assign them to their respective category (accuracy $M = 95.3\%$, SD = 2.66). Each of the objects was used in two different colors (blue and red) and was mirrored to create two different orientations (i.e., the handle pointing to the left or the right). They were presented within a gray rectangle with a size of 306 × 108 pixels. All pictures were analyzed with respect to their size and their luminescence to ensure physical consistency. The manipulations of color and orientation yielded a total of 128 different pictures, which allowed for the presentation of two new pictures in each trial. **Figure 1B** depicts an example of a stimulus screen.

[1] http://www.neurobs.com
[2] www.spss.com

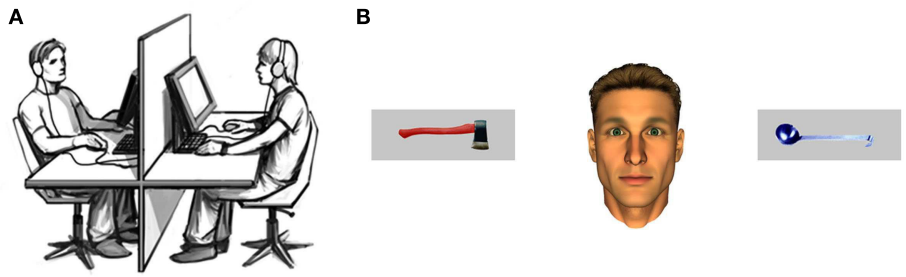

FIGURE 1 | (A) Illustration of the interactive eye-tracking setup with the real participant on one side and the interaction partner – a confederate of the experimenter – on the other (taken from Pfeiffer et al., 2011, p. 2). **(B)** Example trial depicting the male anthropomorphic virtual character and pictures of two real-life objects.

COVER STORY

Participants were led to believe that they would engage in a gaze-based interaction task with another participant and that the interaction would not be vis-à-vis but via virtual characters serving as avatars of their gaze behavior. More specifically, participants were instructed that their eye-movements would be conferred to a virtual character displayed on the screen of their interaction partner. Likewise, the eye-movements of their interaction partner would be visualized by a virtual character displayed on their screen. In fact, however, the interaction partner was a confederate of the experimenter and the virtual character's eye-movements were always controlled by a computer program to ensure full experimental control. Participants were debriefed about this manipulation after the experiment and belief in the cover story was controlled during a post-experiment interview.

PROCEDURE

In the beginning of each experiment the participant and the confederate were seated in front of two eye-tracking devices. Female participants interacted with a female confederate, and male participants with a male confederate, respectively. Subsequently, they received written instructions on the computer screen. A room-divider visually separated both persons. After both of them indicated that they had understood the instructions, the participant's eye-tracker was calibrated. To sustain the cover story, the experimenter pretended to be calibrating the eye-tracker of the interaction partner as well. In addition, during the experiment both persons were asked to wear ear protection so that the participant was not distracted from the task and to make verbal communication impossible.

EXPERIMENT 1A

The first experiment aimed at assessing at which latencies participants experienced gaze reactions – either gaze-following or gaze aversion – of another person as contingent on their own gaze shifts. It consisted of two main conditions: (1) JA trials in which the virtual character followed the participant's gaze and (2) NJA trials in which the virtual character did not follow the participant's gaze but shifted its gaze toward the other object. In both conditions the latency of the virtual character's gaze reactions was varied from 0 to 4000 ms in steps of 400 ms. This yielded eleven sub-conditions which were repeated eight times throughout the experiment, thereby resulting in a total of 176 trials which were presented in a randomized fashion.

Each trial started with an initiation phase in which participants were instructed to fixate the virtual character. Upon fixation two objects appeared to the left and the right of the virtual character. Participants were asked to shift their gaze to one of these objects as quickly as possible and to wait for the reaction of the virtual character. After the character's gaze reaction the scene remained static for another 500 ms before participants had to indicate by button press how strongly related they experienced the gaze reaction of the other to their own gaze shift on a four-item scale (very related – rather related – rather unrelated – very unrelated). Each trial was followed by a short break in which a fixation cross was presented with a latency jittered between 1000 and 2000 ms. The total duration of the experiment was about 25 min.

In this experiment, 30 volunteers participated, out of which 27 (Mean age = 27.63, SD = 6.29, 15 female/12 male) entered the analysis. Two had to be excluded from data analysis because of technical problems and another one due to disbelief in the cover story.

EXPERIMENT 1B

In order to enhance participants' sensitivity to the timing of *gaze-following*, Experiment 1a was repeated without the non-JA condition, that is, the virtual character followed participants' gaze in *all* trials. Participants were instructed that their putative interaction partner was instructed to always look at the same object. As each sub-condition (i.e., reaction latencies from 0 to 4000 ms in steps of 400 ms) was repeated 16 instead of eight times, Experiment 1b did not differ structurally from Experiment 1a.

There were 24 participants in this experiment. Only 21 (Mean age = 23.86, SD = 5.74, 14 female/7 male) were included in the analysis as two had to be excluded due to technical problems and one due to disbelief in the cover story.

EXPERIMENT 2

The aim of this experiment was to assess whether the theoretically proposed processes of JA and SA differ with respect to the interaction dynamics. The experimental design contained a between-subject and a within-subject factor. The within-subject factor was the order of initiation of the interaction sequence (self-initiated vs. other-initiated) and the between-subject factor was task instruction (JA vs. SA). Prior to the experiment, participants were assigned in a randomized but gender-balanced fashion to either a JA or a SA group. In the JA group, participants were instructed to press a response button as soon as *they themselves were aware that both they and their interaction partner directed their attention to the same object*. In the SA condition, participants were asked to press the button as soon as *they were convinced that both of them were aware of each other directing their attention to the same object*. Particular caution was exerted to avoid any explanation that went beyond the descriptions written in italics above and any cues toward the theoretical concepts of JA and SA or related psychological processes.

In both JA and SA groups, the order of initiation of the interaction sequence (i.e., the within-subject factor) was manipulated block-wise. The initiator of a trial is the person who is the first to fixate one of the two objects on the screen. Participants either started with the self-initiated block in the first half of the experiment and then proceeded in the other-initiated block in the second half or vice versa. To avoid sequence effects, participants started with the self- or other-initiated block in an alternating fashion. Each block consisted of 32 trials. In the beginning of each trial two objects were shown for 3000 ms on the left and the right side of the screen so that participants could become acquainted to them and subsequently concentrate on the interaction task. After the acquaintance period the virtual character appeared in the center of the screen. This served as a cue to the initiation of the interaction. Participants were instructed that the establishment

of mutual gaze with the virtual character was a prerequisite for the interaction sequence to start. Depending on the experimental block, there were two ways the interaction period could be initiated. (1) In trials of the self-initiated block participants were told to choose one object by fixating it and the virtual character followed their gaze. (2) In contrast, in trials of the other-initiated block the virtual character commenced the interaction by shifting its gaze to one of the objects. Participants were instructed to follow its gaze. As soon as the first gaze fixation on the virtual character (in the self-initiated condition) or on the chosen object (in the other-initiated condition) was detected, the dynamic interaction period started. When the participant looked at the virtual character, it responded by shifting its gaze to the participant to establish eye contact. When the participant looked back at the object, the virtual character followed his or her gaze. Gaze reactions of the virtual character followed with a latency that was jittered between 400 and 800 ms (i.e., latencies experienced as "natural" for human gaze reactions according to Experiments 1a and 1b). This interaction continued until participants – depending on the group they had been assigned to – indicated the experience of JA or SA (as described above) by pressing a button and thereby ending the current trial.

Overall, 43 participants participated in the study. As three of them were excluded due to technical problems, only 40 of them (Mean age = 24.75, SD = 5.15, 20 female/20 male) were included in the analysis.

RESULTS

EXPERIMENT 1A

The ratings of relatedness of the avatar's gaze reactions are depicted in **Figure 2A**. A two-way ANOVA for repeated-measures with the factors gaze reaction (joint vs. non-joint) and latency (0–4000 ms in steps of 400 ms) showed a main effect of gaze reaction: as expected, gaze-following reactions resulting in JA were experienced as more related to participants' gaze shifts as compared to gaze aversion resulting in NJA, $F(1, 26) = 67.09$, $p < 0.001$. In addition, there was a main effect of latency on participants' ratings of relatedness, $F(5.83, 92.54) = 5.38$, $p = 0.001$ (Greenhouse–Geisser corrected, $\varepsilon = 0.36$, due to a violation of the assumption of sphericity). For both joint and NJA trials, participants rated immediate reactions with a latency of 0 ms as considerably less related to their own gaze shift than reactions with higher latencies. In addition, ratings of relatedness seemed to decrease linearly for latencies greater than 800 ms (see also the "Combined Analysis of Gaze-Following in Experiments 1a and 1b" below). There was no significant interaction between these two factors, $F(6.3, 163.76) = 1.26$, $p = 0.28$.

EXPERIMENT 1B

Figure 2B shows the ratings of relatedness of the avatar's gaze reaction to participants' own gaze shift as a function of the latency of the reaction. A one-way repeated-measures ANOVA revealed that, similar to the results of Experiment 1a, there was a main effect of latency on participants' rating of relatedness of the other's gaze reaction, $F(17.07, 54.87) = 26.78$, $p < 0.001$ (Greenhouse–Geisser corrected, $\varepsilon = 0.27$). This effect was described by a highly significant linear trend, $F(1, 20) = 53.14$, $p < 0.001$, indicating a continuous decrease of relatedness ratings with increasing latency of gaze reactions.

COMBINED ANALYSIS OF GAZE-FOLLOWING IN EXPERIMENTS 1A AND 1B

In a separate set of analyses, we focused only on JA and compared the JA trials from Experiment 1a to Experiment 1b. The crucial difference between these two experiments was that in Experiment 1a the putative interaction partner had an additional option to react and could also avert his/her gaze, whereas in JA, only the latency of the gaze reaction is varied. For better comparability, the joint attention data of Experiment 1a (JA in the context of NJA as another option to act) are plotted together with the data from Experiment 1b (JA only).

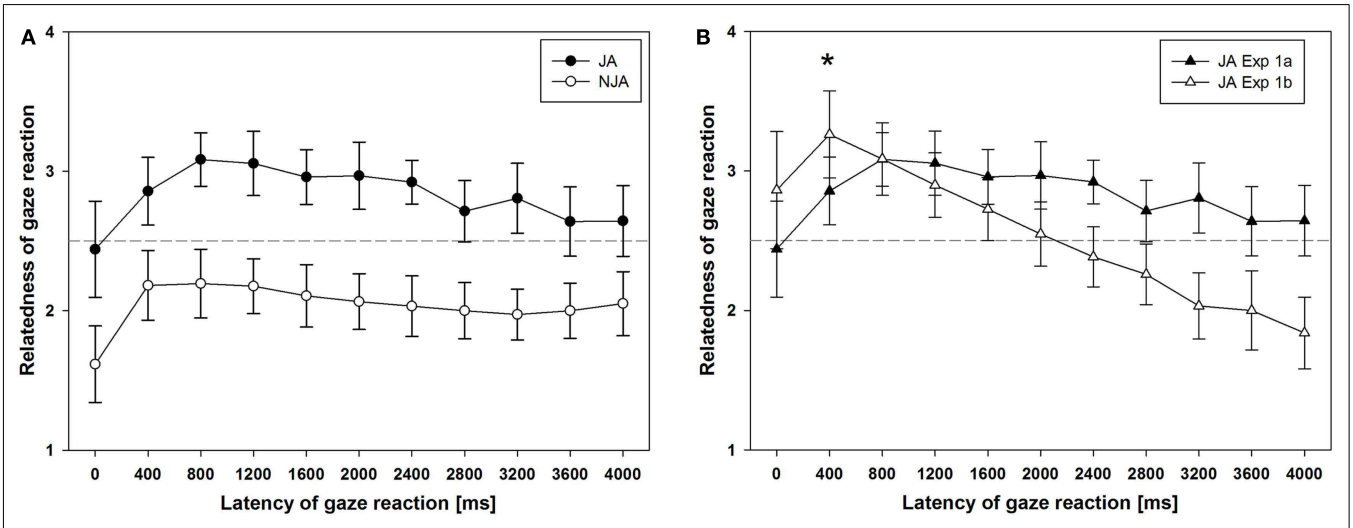

FIGURE 2 | (A) The results from Experiment 1a, in which the interaction partner could either follow the gaze of the participant to engage in joint attention (JA) or avert his/her gaze to the other object to engage in non-joint attention (NJA). **(B)** In Experiment 1b the interaction partner always engaged

Experiment 1b the virtual character would always follow participants' gaze, which participants were informed of during the instruction. In order to assess the influence of a second option to react on the perception of latency of gaze-following, we conducted a two-way repeated-measures ANOVA including only the JA trials from Experiment 1a and all trials from Experiment 1b with experiment as a between-subjects factor. There was a significant interaction between the factors experiment and relatedness rating, $F(4.27, 196.3) = 11.02$, $p < 0.001$ (Greenhouse–Geisser corrected, $\varepsilon = 0.43$). As **Figure 2B** shows, ratings from Experiment 1b (open circles), which consisted only of JA trials, suggest that participants experience gaze-following reactions as most related to their own gaze shift when they follow with a latency of 400 ms ($M = 3.26$, SD = 0.68). In Experiment 1a (filled circles) ratings for gaze reactions with a latency of 400 ms were significantly lower ($M = 2.86$, SD = 0.61), as shown by a t-test for independent samples, $t(46) = -2.16$, $p = 0.038$. Here, visual inspection of data suggests that maximum relatedness ratings were not reached before 800 ms. Furthermore, in Experiment 1b there was a continuous linear decrease of relatedness ratings beginning at 400 ms. This was confirmed by a highly significant linear trend, $F(16.06, 42.67) = 53.14$, $p < 0.001$, which is absent in the data of Experiment 1a, $F(0.47, 17.49) = 0.7$, $p = 0.41$. Taken together, these results suggest that when the interaction partner has no other choice but following participants' gaze, relatedness ratings peak earlier as compared to a context in which the other can either react by gaze-following or by gaze aversion. In addition, participants' are less sensitive to the latency of gaze-following in the context of action alternatives.

EXPERIMENT 2

An independent samples t-test indicated that significantly more gaze shifts were required to reach a situation of shared ($M = 2.55$, SD = 1.26) as compared to JA ($M = 1.23$, SD = 0.35). Furthermore, standard deviations indicate that the inter-individual variance was much higher in SA. This between-subject variance is also depicted in the box plot in **Figure 3A**. Importantly, the establishment of mutual gaze was a prerequisite for the initiation of the interaction to ensure that scan paths always began with a fixation of the virtual character. The increased number of gaze shifts also resulted in significantly longer trial durations in shared ($M = 3886.39$ ms, SD = 1838.91 ms) vs. JA ($M = 2040.11$ ms, SD = 974.64 ms), $t(28.89) = -3.97$, $p < 0.001$, $r = -0.58$. Interestingly, in JA participants showed significantly more gaze shifts in self-initiated trials ($M = 1.41$, SD = 0.68) compared to other-initiated trials ($M = 1.07$, SD = 0.10), $t(19.79) = 2.18$, $p = 0.042$, $r = 0.33$, while there was no such effect of initiation in SA, $t(38) = 0.24$, $p = 0.81$ (see **Figure 3B**), indicating that only the gaze dynamics of JA were influenced by the initiation of the interaction.

DISCUSSION

The present study introduced a novel interactive eye-tracking paradigm suitable to study multiple facets of triadic interactions between two agents and real-world objects in real-time. On a methodological level, this provides an important complement to previous work by our group which has not involved real objects but rather concentrated on the dyadic aspects of gaze-following and JA (Schilbach et al., 2010; Wilms et al., 2010; Pfeiffer et al., 2011). This methodological advancement was used for the empirical investigation of temporal and dynamic aspects of social gaze as a socially salient form of embodied actions with great ecological validity. In Experiments 1a and 1b, participant's sense of agency was measured as a function of both the congruency and latency of another person's gaze reaction. In Experiment 2, differences in gaze dynamics and trial duration resulting in JA and SA were examined. These results provide interesting insights into gaze behavior and the experience of gaze reactions in an ecologically

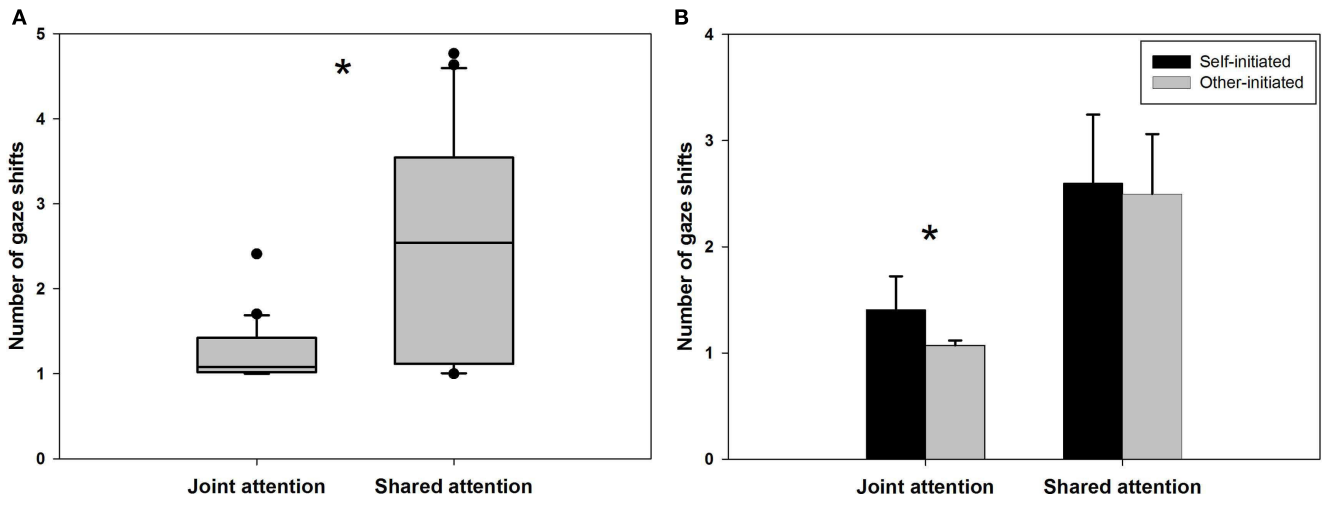

FIGURE 3 | (A) A box plot illustrates the inter-individual variance of the number of gaze shifts before indicating the experience of joint as compared to shared attention. **(B)** Whether participants initiated the gaze-based interaction only affected the number of gaze shifts required to report a state of joint, but not shared attention.

valid but experimentally controllable setting. Conceptual as well as methodological implications are discussed in the following.

EFFECTS OF THE CONGRUENCY OF GAZE REACTIONS

Experiments 1a and 1b investigated how related participants experienced different latencies of gaze reactions to their own gaze behavior by varying these latencies and the congruency of reactions (i.e., gaze-following vs. gaze aversion) systematically. In the following, we suggest that the experience of relatedness can be taken as a measure of the sense of agency (Pacherie, 2012).

It was first predicted that the congruency of the other's gaze reaction (gaze-following vs. gaze aversion) strongly influences participants' sense of agency, as measured by their experience of relatedness. Indeed, results indicated that gaze-following is experienced more strongly related to one's own gaze shifts as compared to gaze aversion. It is highly plausible that this relates to a positive valence that has been associated with gaze-following in comparison to gaze aversion. The literature provides indirect evidence for positive and negative evaluations of gaze-following and gaze aversion, respectively. In a recent study aiming at unraveling the expectations of participants' regarding the behavior of a human interaction partner, we asked participants to interact with a virtual character in a similar interactive eye-tracking setup as in the present study (Pfeiffer et al., 2011). In order to distinguish social from non-social interaction, participants were led to believe that in any given interaction block consisting of a number of gaze trials the virtual character could either be controlled by another person or a computer algorithm. Their task was to decide based on the virtual character's gaze reactions whether they had been interacting with a human or a computer. Unbeknownst to participants, the reactions were always controlled by a computer algorithm to allow full experimental control. Results demonstrated that the proportion of human ratings increased linearly with increasing numbers of gaze-following trials in an interaction block, thereby indicating that in such simple gaze-based interactions, gaze-following and JA are taken as most indicative of true social interaction. This supports the present finding that gaze-following results in an enhanced experience of agency as expressed by higher ratings of self-relatedness.

Another set of studies emphasizes the positive valence of gaze-following in contrast to gaze aversion. A recent study used interactive eye-tracking in an MRI scanner to compare other- and self-initiated situations of JA and NJA and demonstrated a specifically positive valence of self-initiated JA (Schilbach et al., 2010). Results indicated that self-initiated JA correlates with activity in the ventral striatum, a brain region which is a part of the brain's reward system and whose activation has been linked to hedonic experiences (Liu et al., 2007). There is also evidence for negative affective evaluations of gaze aversion. For example, Hietanen et al. (2008) showed in an EEG study that watching pictures of persons averting their gaze leads to avoidance-related neural activity, whereas watching pictures of persons with direct gaze correlated with approach-related signals. Furthermore, persons who avert their gaze are judged as less likeable and attractive as compared to persons exhibiting direct gaze (Mason et al., 2005) and gaze aversion is understood as a non-verbal cue to lying and insincerity (Einav and Hood, 2008; Williams et al., 2009). It is conceivable that the intrinsically rewarding nature of initiating social interaction by leading someone's gaze in combination with the implicitly negative evaluation of averted gaze plays a prominent role in the increased feeling of relatedness for gaze-following as compared to gaze aversion.

THE INFLUENCE OF REACTION LATENCIES AND ACTION POSSIBILITIES ON THE EXPERIENCE OF GAZE REACTIONS

We hypothesized that, while very short latencies might be perceived as coincidental, reactions with long latencies might be experienced as non-contingent upon one's own behavior. Indeed, the most obvious finding was that in all conditions reactions with a latency of 0 ms were experienced as considerably less related than the subsequent latency levels of 400 and 800 ms. This result is plausibly explained by the fact that a certain minimal delay needs to be present until a reaction can be experienced as causally linked to (or launched by) any given preceding action and not just as mere coincidence (Scholl and Tremoulet, 2000). Literature suggests that the natural latency of normal saccades (i.e., not express saccades) to any form of visual displacement on a screen is between 200 and 250 ms (Saslow, 1967; Yang et al., 2002). Although our results do not precisely show at which latencies a reaction is experienced as merely coincidental, it is conceivable that saccadic latencies are implicitly taken into account in participants' ratings of relatedness and that gaze reactions with latencies below 250 ms are therefore considered unrelated. However, further experiments are needed to investigate in detail how latencies of gaze reactions between 0 and 400 ms are experienced.

Notably, however, the experience of different latencies of a gaze-following reaction appears to depend on the other person's options to act. When the other person can choose to follow or to avert her eyes, there is hardly any effect of latency on the experience of relatedness and even reactions with a substantial delay of 4000 ms are experienced as rather related. In contrast, when the other person always engages in gaze-following relatedness ratings decrease linearly starting at a latency of 400 ms. Furthermore, reactions with latencies of more than 2000 ms are experienced as unrelated to one's own gaze shifts – they fall below the dashed line symbolizing a neutral rating in **Figure 2B**, and thereby reach the level of unrelatedness that is associated with NJA.

The effect of the other person's options for action is interesting in that it throws new light on the role of perceived causality for one's sense of agency, which traditionally has to do with predicting the sensory consequences (avatar gaze shift) of self-produced actions (own gaze shift). This means that in a joint context, whereas my sensorimotor cues with respect to my own action remain identical to non-joint situations, I perceive the consequences of my actions *in the actions of the other person*. Therefore, the nature of the other person's behavior will have a bearing on my experience of self-agency. In particular, as Pacherie (2012) notes, the strength of the sense of agency is related to how well our predictions regarding another person's reaction to our own actions match with the actual reaction. This is specifically true in small-scale interactions – as in our experiments – in which every aspect

of the interactors' behavior is accessible. Rather than investigating sense of agency in an all-or-none fashion, we therefore interpreted participants' ratings of relatedness of the other's gaze reaction as a measure of how strongly they experienced agency in a given gaze trial.

Adopting this view of agency, the results of experiments 1a and 1b could reflect the role of perceived causality for one's sense of agency. Haggard et al. (2002) have suggested that sense of agency depends crucially on the intentionality of the agent and found that it decreases with increasing action-outcome delays, as it does in Experiment 1b, and to a lesser degree in Experiment 1a. Subsequent research has shown that not only intentionality, but also *perceived* causality is crucial for the sense of agency. Buehner and Humphreys (2009) found that, when keeping action-outcome constant, given a strong perceived causal link, intentional binding was preserved at action – outcome delays of up to 4 s, as in Experiment 1a. However, there is a less persistent sense of agency in Experiment 1b although the actual causal link is stronger due to the avatar always following my gaze. This could mean that perceived causality is less important for my sense of agency in an interactive context. More plausibly, it could be that in an interactive context, since I am dealing with another agent, the evaluation of my own actions as causally efficacious is only meaningful *when I know that the other has different options for action*. Put otherwise, if I have to evaluate my own sense of agency, *given* that the effect is observed in the behavior of *another* agent, my judgment could be influenced crucially by the sense of agency I am able to attribute to the other (as suggested in Schilbach et al., in press). Further research is needed to look at the interdependency of one's sense of agency for self and other in interaction, but the data from the first experiment show that there is a difference between how sense of agency is experienced in social as compared to non-social situations.

DIFFERENCES IN GAZE DYNAMICS BETWEEN JOINT AND SHARED ATTENTION

In Experiment 2, the dynamics of gaze behavior in situations of JA and SA were assessed while making use of the temporal parameters uncovered in Experiment 1b. As described in the introduction, the necessary criteria for *joint attention* require only one of the interaction partners to be aware of the joint focus of attention. *Shared attention*, however, warrants *both* gazer and gaze-follower to be simultaneously aware of focusing on the same object *and* on each other's awareness of focusing on the same object (Emery, 2000). Results clearly indicate that participants required a significantly higher number of gaze shifts between objects and the virtual character in order to establish SA as compared to JA. As a consequence of this, trial length was considerably longer. JA required only slightly more than one gaze shift on average and is reached significantly earlier in self- vs. other-initiated trials. This indicates that participants were able to make inferences about the emergence of JA by focusing on the object and seemingly observing their partner's gaze reaction at the same time. Due to the impossibility of fixating two spatially separated objects simultaneously, these data demonstrate that a peripheral and quick recognition of the other's gaze reaction is sufficient for the establishment of JA. In contrast to SA, the establishment of JA happens rapidly and is characterized by considerably less inter-individual invariance (see **Figure 3A**). This suggests that JA is characterized by the mere detection of the other's focus of attention, thereby possibly representing a visual detection task rather than a mentalizing task. Unfortunately, it is not directly possible to compare reaction times between the present results and findings on visual detection. Previous studies have not used interactive settings but concentrated on the detection of objects in real-world scenes (Biederman, 1972) or on the detection of gaze direction in static displays (Franck et al., 1998). Using interactive eye-tracking, however, the link between JA and visual detection could now be assessed specifically.

In contrast, such an observation of the other's gaze behavior "out of the corner of the eyes" appears to be insufficient for a reliable identification of a situation of SA. It has previously been argued that SA might be characterized by an increased level of interactivity (Staudte and Crocker, 2011). According to Kaplan and Hafner (2006), true SA requires a monitoring and understanding of the intentions of the other in a coordinated interaction process and is only reached when "both agents are aware of this coordination of "perspectives" toward the world" (Kaplan and Hafner, 2006, p. 145). The increased number of gaze shifts between the virtual character's face and the object and the correlated increase in trial length are indicative of such a coordinated interaction aimed at an alignment of intentions. Determining whether another person is aware of the object jointly focused upon as well as of "us" being aware of us being aware requires thinking about the other's mental states. This is reflected by the dynamics of gaze behavior which exceed the simple detection of a gaze shift to a joint focus of attention. In the vast majority of trials in the JA condition there is not a single look back to the virtual character's face, while this is practically always the case in the SA condition (**Figure 3**): participants have to re-establish eye contact at least once before they indicate to experience SA. It has recently also been shown in an interaction task within a minimalist virtual environment that higher complexity and reciprocity in the dynamics of a *tactile* interaction leads to the experience of interacting with another human agent (Auvray et al., 2009). The experience of non-verbal social interaction therefore more generally seems to hinge upon certain elaborate dynamics between actions and reactions.

A final observation refers to the substantial inter-individual variance in the number of gaze shifts participants exhibit before indicating the experience of SA (cf. **Figure 3A**). This connotes that gaze behavior as an embodied correlate of mentalizing is subject to greater inter-individual differences as compared to gaze behavior in a visual detection task. Literature suggests that inter-individual differences in personality traits and behavioral dispositions strongly influence the performance in different types of mentalizing tasks, i.e., tasks that require reasoning about other persons' mental states. For example, self-reported measures of empathy (Baron-Cohen and Wheelwright, 2004) or of the drive to do things systematically (i.e., systemizing, Baron-Cohen et al., 2003) as well as the personality trait of agreeableness (for a detailed discussion, see Nettle and Liddle, 2008) have been shown to affect mentalizing in a variety of tasks. More studies are

required in order to determine which personality traits or behavioral dispositions result in the observed variance of gaze patterns in SA.

Taken together, the findings reported in this paper can be taken as a first fine-grained description of the temporal and spatial dynamics of social gaze in triadic interactions and their influence on our sense of agency and awareness of the mental states of others. Further assessment of the underlying mental processes is required to understand how manipulations of these aspects change our experience of a social interaction and our perception of the interaction partner.

OUTLOOK

Interactive eye-tracking paradigms incorporating virtual characters have proven specifically useful for the study of social interaction face-to-face and in real-time (Schilbach et al., in press). One major asset of such studies is that the results can be immediately fed back into novel designs with even greater ecological validity. This can stimulate the development for therapeutic tools to learn or improve non-verbal communication in autism spectrum disorders. These are characterized by impairments of the ability to interact with others, as well as by a specific deficiency in reading information from the eye region and interpreting gaze cues (Senju and Johnson, 2009). For example, autistic persons have problems engaging in JA – this is most apparent for the initiation of JA, although responding to another person's bid for JA can also be problematic (Mundy and Newell, 2007). In a recent report on attempts to teach autistic children to initiate and respond to bids of JA, they were required to engage in triadic interactions with an instructor and different kinds of toys (Taylor and Hoch, 2008). As this setting made eye contact difficult, JA was initiated by the instructor by pointing at an object instead of gazing at it. In the condition in which the children were supposed to initiate JA, they were prompted verbally to do so and explicitly told how to do it. A gaze-contingent display would be advantageous here for several reasons: first of all, the interaction with an avatar would be less distressing for autistic persons than real social interaction. Especially in the beginning of a training program this might be beneficial. Secondly, the training program could be designed in a highly structured manner. Features of the avatar's gaze behavior such as timing, gaze direction, or the length of direct gaze could be varied systematically while other facial features can be kept constant in order to prevent sensory overload. Thirdly, the simultaneous recording of eye-movements can be used to analyze scan paths in order to detect difficulties or peculiarities in the participant's gaze behavior. Furthermore, using interactive eye-tracking allows changing the avatar's reactions depending on the participant's gaze behavior in real-time. Lastly, a virtual setting provides more options to highlight and manipulate objects, prompt certain actions, or deliver reinforcement for correct behavior.

Very recently, first attempts have been made to design gaze-contingent virtual reality applications (Bellani et al., 2011; Lahiri et al., 2011). Lahiri et al. (2011) designed a virtual reality application for autistic adolescents in which they are required to interact with a realistically designed virtual classmate. Their task was to make this classmate as comfortable as possible by their behavior. They were positively reinforced the more they looked at the eyes of the character or followed their movements to an object on the screen. A gaze-contingent algorithm inspired by the one invented by Wilms et al. (2010) was used to detect fixations within predefined regions of interest (i.e., eyes, face, object) and to determine the kind of reinforcement depending on when and how long these regions were fixated. This provides a very interesting example for an implicit training of non-verbal social skills using a gaze-sensitive virtual environment. Although this approach is promising, therapeutic tools still have difficulties providing the avatars with realistic gaze behavior (Bellani et al., 2011). Although clearly more work is needed, results from the present study could potentially be incorporated into virtual therapeutic tools.

CONCLUSION

A thorough exploration and understanding of the parameters of social gaze is crucial for the investigation and understanding of social interactions in gaze-contingent paradigms (Wilms et al., 2010; Bayliss et al., 2012; Grynszpan et al., 2012) and for the formulation of hypotheses regarding people's gaze behavior in online interaction (Neider et al., 2010; Dale et al., 2011). In addition, recent advances have been made to the development of dual eye-tracking setups which allow for investigating the gaze behavior of two participants interacting and collaborating in a shared virtual environment (Carletta et al., 2010). Although this approach is very promising, the design of tasks allowing for an assessment of interaction dynamics while controlling variables affecting the interaction still remains a challenge. Before true interaction without simulated others can be investigated, the use of interactive eye-tracking paradigms provides an important tool to study social gaze behavior in persons who experience being engaged and being responded to in an interaction.

ACKNOWLEDGMENTS

This study was partially supported by a grant of the Köln Fortune Program of the Medical Faculty at the University of Cologne to Leonhard Schilbach and by a grant "Other Minds" of the German Ministry of Research and Education to Kai Vogeley. The authors would like to thank Stephanie Alexius and Leonhard Engels for their assistance in data collection.

REFERENCES

Argyle, M., and Cook, M. (1976). *Gaze and Mutual Gaze*. Cambridge, UK: Cambridge University Press.

Auvray, M., Lenay, C., and Stewart, J. (2009). Perceptual interactions in a minimalist virtual environment. *New Ideas Psychol.* 27, 32–47.

Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Pers. Soc. Psychol. Bull.* 29, 819–833.

Bandura, A. (1989). Human agency in social cognitive theory. *Am. Psychol.* 44, 1175–1184.

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., and Wheelwright, S. (2003). The systemizing quotient: an investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 361–374.

Baron-Cohen, S., and Wheelwright, S. (2004). The empathy quotient: an investigation of adults

with Asperger syndrome or high functioning autism, and normal sex differences. *J. Autism Dev. Disord.* 34, 163–175.

Bayliss, A. P., Frischen, A., Fenske, M. J., and Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition* 104, 644–653.

Bayliss, A. P., Murphy, E., Naughtin, C. K., Kritikos, A., Schilbach, L., and Becker, S. I. (2012). "Gaze Leading": initiating simulated joint attention influences eye movements and choice behavior. *J. Exp. Psychol. Gen.* [Epub ahead of print].

Bayliss, A. P., Paul, M. A., Cannon, P. R., and Tipper, S. P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. *Psychon. Bull. Rev.* 13, 1061–1066.

Becchio, C., Sartori, L., and Castiello, U. (2010). Toward you: the social side of actions. *Curr. Dir. Psychol. Sci.* 19, 183–188.

Bellani, M., Fornasari, L., Chittaro, L., and Brambilla, P. (2011). Virtual reality in autism: state of the art. *Epidemiol. Psychiatr. Sci.* 20, 235–238.

Bente, G., Rüggenberg, S., Krämer, N. C., and Eschenburg, F. (2008). Avatar-mediated networking: increasing social presence and interpersonal trust in net-based collaborations. *Hum. Commun. Res.* 34, 287–318.

Biederman, I. (1972). Perceiving real-world scenes. *Science* 177, 77–80.

Bohil, C. J., Alicea, B., and Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nat. Rev. Neurosci.* 12, 752–762.

Buehner, M. J., and Humphreys, G. R. (2009). Causal binding of actions to their effects. *Psychol. Sci.* 20, 1221–1228.

Carletta, J., Hill, R. L., Nicol, C., Taylor, T., de Ruiter, J. P., and Bard, E. G. (2010). Eyetracking for two-person tasks with manipulation of a virtual world. *Behav. Res. Methods* 42, 254–265.

Carpenter, M., Nagell, K., and Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* 63, i–vi, 1–143.

Clark, H. H. (1996). *Using Language*, 2nd Edn. Cambridge: Cambridge University Press.

Dale, R., Kirkham, N. Z., and Richardson, D. C. (2011). The dynamics of reference and shared visual attention. *Front. Psychol.* 2:355. doi:10.3389/fpsyg.2011.00355

de Vignemont, F., and Fourneret, P. (2004). The sense of agency: a philosophical and empirical review of the "Who" system. *Conscious. Cogn.* 13, 1–19.

Einav, S., and Hood, B. M. (2008). Tell-tale eyes: children's attribution of gaze aversion as a lying cue. *Dev. Psychol.* 44, 1655–1667.

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neurosci. Biobehav. Rev.* 24, 581–604.

Franck, N., Daprati, E., Michel, F., Saoud, M., Daléry, J., Marie-Cardine, M., et al. (1998). Gaze discrimination is unimpaired in schizophrenia. *Psychiatry Res.* 81, 67–75.

Frischen, A., Bayliss, A. P., and Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychol. Bull.* 133, 694–724.

Frith, C. D., and Frith, U. (2006). The neural basis of mentalizing. *Neuron* 50, 531–534.

Gibson, J. J., and Pick, A. D. (1963). Perception of another person's looking behavior. *Am. J. Psychol.* 76, 386–394.

Gopnik, A., Slaughter, V., and Meltzoff, A. (1994). "Changing your views: how understanding visual perception can lead to a new theory of the mind," in *Children's Early Understanding of Mind: Origins and Development*, eds C. Lewis and P. Mitchell (Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.), 157–181.

Grynszpan, O., Simonin, J., Martin, J.-C., and Nadel, J. (2012). Investigating social gaze as an action-perception online performance. *Front. Hum. Neurosci.* 6:94. doi:10.3389/fnhum.2012.00094

Haggard, P., Clark, S., and Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nat. Neurosci.* 5, 382–385.

Hietanen, J. K., Leppänen, J. M., Peltola, M. J., Linna-aho, K., and Ruuhiala, H. J. (2008). Seeing direct and averted gaze activates the approach-avoidance motivational brain systems. *Neuropsychologia* 46, 2423–2430.

Kaplan, F., and Hafner, V. V. (2006). The challenges of joint attention. *Interact. Stud.* 7, 135–169.

Lahiri, U., Warren, Z., and Sarkar, N. (2011). Design of a gaze-sensitive virtual social interactive system for children with autism. *IEEE Trans. Neural Syst. Rehabil. Eng.* 19, 443–452.

Liu, X., Powell, D. K., Wang, H., Gold, B. T., Corbly, C. R., and Joseph, J. E. (2007). Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. *J. Neurosci.* 27, 4587–4597.

Loomis, J. M., Blascovich, J. J., and Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behav. Res. Methods Instrum. Comput.* 31, 557–564.

Mason, M. F., Tatkow, E. P., and Macrae, C. N. (2005). The look of love: gaze shifts and person perception. *Psychol. Sci.* 16, 236–239.

Moll, H., and Tomasello, M. (2007). Cooperation and human cognition: the Vygotskian intelligence hypothesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 639–648.

Mundy, P., and Newell, L. (2007). Attention, joint attention, and Social Cognition. *Curr. Dir. Psychol. Sci.* 16, 269–274.

Neider, M. B., Chen, X., Dickinson, C. A., Brennan, S. E., and Zelinsky, G. J. (2010). Coordinating spatial referencing using shared gaze. *Psychon. Bull. Rev.* 17, 718–724.

Nettle, D., and Liddle, B. (2008). Agreeableness is related to social-cognitive, but not social-perceptual, theory of mind. *Eur. J. Pers.* 22, 323–335.

Pacherie, E. (2012). "The phenomenology of joint action: self-agency vs. joint-agency," in *Joint Attention: New Developments*, ed. A. Seemann (Cambridge, MA: MIT Press), 343–389.

Pfeiffer, U. J., Timmermans, B., Bente, G., Vogeley, K., and Schilbach, L. (2011). A non-verbal turing test: differentiating mind from machine in gaze-based social interaction. *PLoS ONE* 6:e27591. doi:10.1371/journal.pone.0027591

Povinelli, D. J., and Eddy, T. J. (1996). Chimpanzees: joint visual attention. *Psychol. Sci.* 7, 129–135.

Redcay, E., Dodell-Feder, D., Pearrow, M. J., Mavros, P. L., Kleiner, M., Gabrieli, J. D. E., et al. (2010). Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. *Neuroimage* 50, 1639–1647.

Saslow, M. G. (1967). Effects of components of displacement-step stimuli upon latency for saccadic eye movement. *J. Opt. Soc. Am.* 57, 1024–1029.

Schilbach, L., Eickhoff, S. B., Cieslik, E. C., Kuzmanovic, B., and Vogeley, K. (2011). Shall we do this together Social gaze influences action control in a comparison group, but not in individuals with high-functioning autism. *Autism* 16, 151–162.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., et al. (in press). Toward a second-person neuroscience. *Behav. Brain Sci*.

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., et al. (2010). Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. *J. Cogn. Neurosci.* 22, 2702–2715.

Scholl, B., and Tremoulet, P. (2000). Perceptual causality and animacy. *Trends Cogn. Sci. (Regul. Ed.)* 4, 299–309.

Senju, A., and Johnson, M. H. (2009). Atypical eye contact in autism: models, mechanisms and development. *Neurosci. Biobehav. Rev.* 33, 1204–1214.

Shepherd, S. V. (2010). Following gaze: gaze-following behavior as a window into social cognition. *Front. Integr. Neurosci.* 4:5. doi:10.3389/fnint.2010.00005

Staudte, M., and Crocker, M. W. (2011). Investigating joint attention mechanisms through spoken human-robot interaction. *Cognition* 120, 268–291.

Stern, D. N. (1974). The goal and structure of mother-infant play. *J. Am. Acad. Child Psychiatry* 13, 402–421.

Synofzik, M., Vosgerau, G., and Newen, A. (2008). Beyond the comparator model: a multifactorial two-step account of agency. *Conscious. Cogn.* 17, 219–239.

Tarr, M. J., and Warren, W. H. (2002). Virtual reality in behavioral neuroscience and beyond. *Nat. Neurosci.* 5, 1089–1092.

Taylor, B. A., and Hoch, H. (2008). Teaching children with autism to respond to and initiate bids for joint attention. *J. Appl. Behav. Anal.* 41, 377–391.

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–691; discussion 691–735.

van der Weiden, A., Veling, H., and Aarts, H. (2010). When observing gaze shifts of others enhances object desirability. *Emotion* 10, 939–943.

Vogeley, K., and Bente, G. (2010). "Artificial humans": psychology and neuroscience perspectives on embodiment and nonverbal communication. *Neural. Netw.* 23, 1077–1090.

von der Pütten, A. M., Krämer, N. C., Gratch, J., and Kang, S.-H. (2010). "It doesn't matter what you are!" Explaining social effects of agents and avatars. *Comput. Human Behav.* 26, 1641–1650.

Williams, J. A., Burns, E. L., and Harmon, E. A. (2009). Insincere utterances and gaze: eye contact during sarcastic statements. *Percept. Mot. Skills* 108, 565–572.

Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G. R., and Vogeley, K. (2010). It's in your eyes – using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Soc. Cogn. Affect. Neurosci.* 5, 98–107.

Yang, Q., Bucci, M. P., and Kapoula, Z. (2002). The latency of saccades, vergence, and combined eye movements in children and in adults. *Invest. Ophthalmol. Vis. Sci.* 43, 2939–2949.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 July 2012; accepted: 13 November 2012; published online: 03 December 2012.

Citation: Pfeiffer UJ, Schilbach L, Jording M, Timmermans B, Bente G and Vogeley K (2012) Eyes on the mind: investigating the influence of gaze dynamics on the perception of others in real-time social interaction. Front. Psychology ***3****:537. doi: 10.3389/fpsyg.2012.00537*

This article was submitted to Frontiers in Cognitive Science, a specialty of Frontiers in Psychology.

OPEN ACCESS Freely available online

PLoS one

A Non-Verbal Turing Test: Differentiating Mind from Machine in Gaze-Based Social Interaction

Ulrich J. Pfeiffer[1,3]*, Bert Timmermans[1], Gary Bente[2], Kai Vogeley[1,3], Leonhard Schilbach[1,4]

1 Neuroimaging Group, Department of Psychiatry and Psychotherapy, University Hospital Cologne, Cologne, Germany, **2** Department of Media and Social Psychology, Faculty of Human Sciences, University of Cologne, Cologne, Germany, **3** Institute of Neuroscience and Medicine – Cognitive Neurology (INM3), Research Center Jülich, Jülich, Germany, **4** Max-Planck-Institute for Neurological Research, Cologne, Germany

Abstract

In social interaction, gaze behavior provides important signals that have a significant impact on our perception of others. Previous investigations, however, have relied on paradigms in which participants are passive observers of other persons' gazes and do not adjust their gaze behavior as is the case in real-life social encounters. We used an interactive eye-tracking paradigm that allows participants to interact with an anthropomorphic virtual character whose gaze behavior is responsive to where the participant looks on the stimulus screen in real time. The character's gaze reactions were systematically varied along a continuum from a maximal probability of gaze aversion to a maximal probability of gaze-following during brief interactions, thereby varying contingency and congruency of the reactions. We investigated how these variations influenced whether participants believed that the character was controlled by another person (i.e., a confederate) or a computer program. In a series of experiments, the human confederate was either introduced as *naïve* to the task, *cooperative*, or *competitive*. Results demonstrate that the ascription of humanness increases with higher congruency of gaze reactions when participants are interacting with a naïve partner. In contrast, humanness ascription is driven by the degree of contingency irrespective of congruency when the confederate was introduced as cooperative. Conversely, during interaction with a competitive confederate, judgments were neither based on congruency nor on contingency. These results offer important insights into what renders the experience of an interaction truly social: Humans appear to have a default expectation of reciprocation that can be influenced drastically by the presumed disposition of the interactor to either cooperate or compete.

Citation: Pfeiffer UJ, Timmermans B, Bente G, Vogeley K, Schilbach L (2011) A Non-Verbal Turing Test: Differentiating Mind from Machine in Gaze-Based Social Interaction. PLoS ONE 6(11): e27591. doi:10.1371/journal.pone.0027591

Editor: Mark W. Greenlee, University of Regensburg, Germany

Received June 2, 2011; **Accepted** October 20, 2011; **Published** November 9, 2011

Funding: This research has been funded by a grant of the Köln Fortune Program of the Medical Faculty to Leonhard Schilbach (http://www.medfak.uni-koeln.de/195.html). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ulrich.pfeiffer@uk-koeln.de

Introduction

In the last decades, considerable knowledge has been acquired about how we perceive other persons, how we interpret their non-verbal behavior, and how we 'read' their minds. However, most experimental paradigms used to this end have relied on testing individuals in isolation. Thus, social interaction is investigated *without* interaction ('offline' social cognition), seemingly reflecting the view that social cognition can be sufficiently understood by investigating what a single person thinks or believes [1]. In recent years, this cognitivist and individualist approach to social cognition has been subject to criticism as it fails to incorporate the interaction process in itself, i.e. the embodiment of agents in an interaction, and the situated nature of social interaction ('online' social cognition, [2]). Instead, enactive accounts of social cognition have gained popularity and suggest to investigate interaction partners in true dyadic interactions [1,3–5]. These accounts are based on the propositions that i) perception and action are inseparable from each other, and that ii) meaning emerges from the active exploration of and coupling with the environment.

One major reason for the scarcity of truly interactive studies in social cognition research might be the complexity of studying complex social interaction processes involving the exchange of subtle and transient cues under standardized laboratory conditions. However rich everyday social interactions present themselves, it is of great importance that the bandwidth of the interaction is restricted substantially in order to study core processes of interaction whilst maintaining acceptable levels of experimental control. Keeping this in mind, any endeavor of assessing real social interaction in fact faces two major challenges. First, an experimentally controllable domain of social cues needs to be identified. Second, a task that reliably separates and contrasts social and non-social interaction must be established.

The first challenge can be met by starting from a subset of communicative cues, which have high explanatory value for social cognitive processes and exchange in social encounters and are at the same time objectively measurable and controllable in an experimental setting. Such a cue system is ideally represented by human gaze. Gaze behavior has long been demonstrated to provide a highly informative window into social cognition [6,7]. Here, an important aspect of social interaction is the ability to follow another person's gaze and share a perceptual experience with someone else, thereby engaging in triadic relations between self, other, and the environment in joint attention [8]. Joint attention is believed to be crucial for an understanding of other minds [9]. An essential distinction has been made with respect to

the person who is initiating joint attention and who is responding to bids of joint attention [10]. In line with observations from non-typically developing humans and research in non-human species, Moll and Tomasello [11] argue that the natural motivation to engage others in triadic interactions represents a uniquely human cognitive factor which might ultimately foster the development of a shared reality [12]. In addition, as the act of looking is both a source of stimulation and a response, perception and action are inseparable in this channel of non-verbal behavior and can hence be subject to tight experimental control [13].

A powerful paradigm to analyze social gaze in a truly interactive way has been introduced recently [14] using interactive eye-tracking and gaze-contingent eye movement simulation. This setup allows to track a person's gaze on a stimulus screen and to control the gaze behavior of an anthropomorphic virtual character [15] dependent on the current gaze position. For the first time, this permits the exploration of gaze-based social interaction in an experimentally controllable way. In an initial study employing this interactive eye-tracking setup in a functional magnetic resonance imaging (fMRI) environment, it could be shown that self-initiated joint attention, i.e. making the virtual character follow one's own gaze, recruits reward-related neurocircuitry consistent with the above described idea of an intrinsic motivation to jointly attend to aspects of the environment [16].

Based on this paradigm, we have developed a gaze-based version of what is known as the "Turing test" in order to study which parameters of gaze-based interactions influence humanness ratings of the virtual character. The Turing test was proposed by the British mathematician Alan Turing in order to address the question whether machines can think, i.e., whether or under which circumstances humans would ascribe human-like intelligence to machines. In order to address this question he suggested various experiments, one of which later became known as the standard Turing test. In this test, a human participant engages in verbal conversation via a computer screen with another human and a computer placed in separate rooms via a computer screen and has to judge with whom he is interacting [17]. If the participant cannot reliably distinguish between the human and the computer conversation partner, the machine is said to have passed the test. The rationale of this paradigm was used in our study to investigate humanness ascriptions during interaction.

For this purpose, we created a gaze-based version of the Turing Test, which in the following will be referred to as the "non-verbal Turing test". In this test participants engage in the ascription of human agency during social interaction, which will be referred to as "ascription of humanness" throughout this article. They have to judge whether they interact with a real human or a computer based on the gaze behavior displayed by an anthropomorphic virtual character in response to their own gaze behavior (see Fig. 1a), while in fact the latter is always the case and the putative other participant is a confederate of the experimenter. Each interaction between participant and agent consisted of six events, during each of which the virtual character would either follow the participant's gaze toward an object that was also shown on the screen or look away from that object (see Fig. 1b). The experimental manipulation consisted in the systematic variation of the number of gaze-following reactions from zero (i.e. character always looking in the opposite direction) to six (i.e. character always following) out of six possible times. In a between-subject design, we also addressed the influence of prior knowledge about the putative interactor's behavioral predisposition in order to model different social contexts. To this end, we introduced the interactor as either *naïve* to the task, *cooperative*, or *competitive*.

Based on the literature we hypothesized three distinct outcomes in the different conditions: *(1) Congruency-based judgment in naïve interaction:* The significance of self-initiated joint attention in social cognition has been highlighted above. Particularly the data by Schilbach et al [16] suggest a motivational aspect of initiating joint attention that is reflected both on the neural and the behavioral level. This might be taken to suggest that humanness ascription should increase with increasing congruency of gaze behavior, i.e. that the experience of interacting with another person increases with the degree of gaze-following when nothing else is known about this person. *(2) Contingency-based judgment in cooperative interaction:* In definitions of cooperation, particular emphasis is put on the necessity of coordination between the cooperative interactors [18]. Therefore, we hypothesize that any form of coordinated reactions could be taken as indicative of a human interaction partner. Importantly, not only maximal gaze-following but also maximal gaze aversion is a highly coordinated interaction pattern as both patterns are maximally contingent upon the participant's gaze. The difference with respect to the participant's gaze is that one pattern is congruent and the other is incongruent. Hence, if coordination played a greater role in humanness ascription when encountering a cooperative interactor, contingent rather than merely congruent reactions should inform participant's

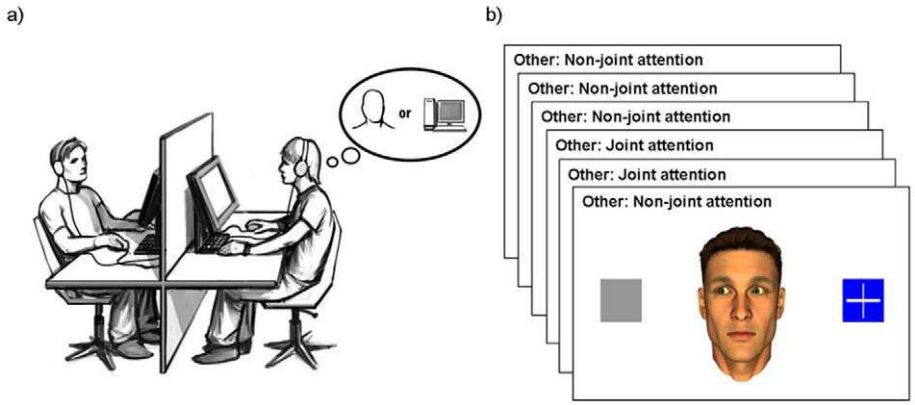

Figure 1. The non-verbal Turing test. *(a)* Set-up of the experiment with a volunteer participating in the study on the right and a confederate of the experimenter acting as a putative interaction partner on the left. *(b)* One exemplar interaction block of the experiment consisting of six interaction events.
doi:10.1371/journal.pone.0027591.g001

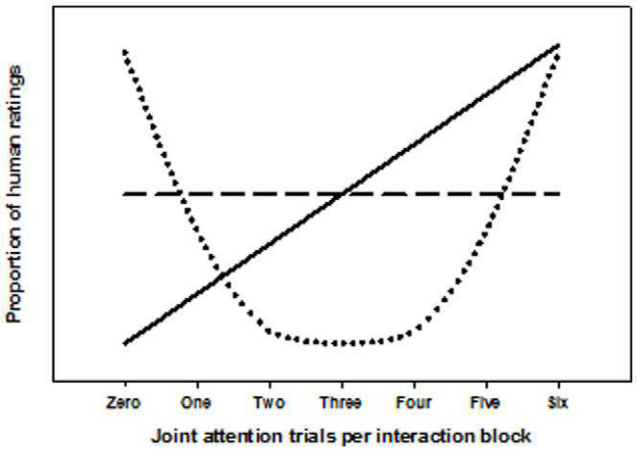

Figure 2. Hypotheses of humanness ascription under changing situational demands are depicted here as simple models. (1) Naive interaction: The ascription of humanness is based on maximally congruent reactions (solid line). (2) Cooperative interaction: The ascription of humanness is based on the mere contingency of reactions (dotted line). (3) Competitive interaction: The ascription of humanness is neither based on congruency nor on contingency of gaze reactions (dashed line).
doi:10.1371/journal.pone.0027591.g002

judgments. *(3) Unpredictability-based judgment in competitive interaction:* In the light of the hypotheses on how humanness is ascribed in situations with a naïve or a cooperative interactor, it might be anticipated that participants would expect a competitive person to avoid any patterned response and hence will not interpret any form of congruency or contingency as indicative of a competitive interactor. Figure 2 provides an illustration of these hypotheses.

Methods

Participants

In total, 128 healthy male and female volunteers aged 19 to 42 years (mean age $=26.72\pm5.31$), with no record of neurologic or psychiatric illnesses participated in the study. They were recruited using an internet-based system [19]. All participants were naïve with respect to the task and to the scientific purpose of the study and were equally compensated for their participation (10 Euro/hour). In the beginning of the study participants were asked to sign a written consent form in which they approved that participation is voluntary and that data are used in an anonymized fashion for statistical analysis and scientific publication. The study strictly followed the WMA Declaration of Helsinki (Ethical Principles for Medical Research Involving Human Subjects) and was presented to and approved by the ethics committee of the medical faculty of the University of Cologne, Germany.

Setup and Materials

We made use of a recently developed interactive eye-tracking paradigm [14]. This method allows participants to interact with an anthropomorphic virtual character by means of their eye-movements. In order to detect participants' eye-movements we used a high resolution eye-tracking system with a digitization rate of 50 Hz and an accuracy of 0.5° (TobiiTM T1750 Eye-Tracker, Tobii Technology AB, Sweden). Participants were seated at a distance of 80 cm in front of the device. Stimuli were presented on the 17" TFT screen of the eye-tracking device with screen resolution set to 1024 by 768 pixels. The viewing angle was 32×24 degrees for the whole screen. A PC with a dual-core processor and a GeForce 2 MX graphics board controlled the output of the eye-tracker as well as stimulus presentation at a frame rate of 100 Hz. Via a fast network connection gaze position updates were transferred to dedicated gaze extraction software (ClearviewTM, Tobii Technology AB, Sweden) which produced real-time gaze position output. This was made available to and used by the Presentation software (PresentationTM, http://www.neurobs.com) to control stimuli in a gaze-contingent manner.

Task

The interaction was organized in interaction blocks of six events each (Fig. 2b). Each of these events had the following order: Participants first had to look at the virtual character. Once the program had detected a fixation of the virtual character two grey squares appeared on the left and the right side of the screen (see [14] for details on the gaze processing algorithm). Participants subsequently had to choose one of the squares by fixating it. Upon fixation the chosen square changed its color from grey to blue to provide feedback about successful gaze detection for the participant. Participants were told that their first gaze to one of the squares (but not the color change) was transferred to the screen of the eye-tracking device of the other participant in real time and that they would see the other participant's response to this as visualized by the eyes of the virtual character visible on their stimulus screen.

As part of the "cover story", participants were told that in a given interaction block the eye-movements of the virtual character could either be controlled by the partner or by a computer program. After each block, the participant's task was to judge whether they had been interacting with the human partner or with the computer program. In actual fact, the other person was a confederate of the experimenter and the eye-movements of the virtual character were always controlled by the computer algorithm. Interaction blocks consisted of six interaction trials, thus allowing for a systematic manipulation of the virtual character's gaze-following or gaze aversion behavior from zero to six out of six (0/6 to 6/6) possible times. Gaze-following thereby constituted a joint attention event, whereas gaze aversion constituted a non-joint attention event. Overall, this resulted in seven conditions (0/6, 1/6, 2/6, 3/6, 4/6, 5/6, 6/6) each of which was repeated eight times in a fully randomized fashion during the course of the experiment. The latency of the virtual character's gaze reaction was jittered between 350 and 600 milliseconds. This resulted in gaze latencies that have previously been found to appear natural to participants (unpublished data). Joint and non-joint attention events were distributed randomly within each interaction block. At the end of each block participants were asked to indicate via button press whether they had been interacting with the other person or the computer program.

Procedure

At the beginning, participants were seated at a distance of about 80 cm from the eye-tracking device. Instructions were provided in a standardized manner on the screen. Participants were informed that during the experiment they would be asked to engage in interaction with a virtual character presented on a computer screen in front of them by looking at the character and by looking at objects also visible on the screen. After the participant was briefed (see descriptions of experiments 1 – 5 for details), the confederate (in the following referred to as the "interactor"), who was said to be instructed simultaneously by a second experimenter in a different room, was brought into the testing room and seated in front of the second eye-tracking device. The two persons were placed about 4 meters apart from each other and were visually

separated by a room-divider. The experimenter then engaged in a brief, scripted conversation with the interactor, thereby repeating some of the instructions to make the cover story believable for the actual participant. Before the experiment started, the participants' sitting position in front of the eye-tracker was optimized and the eye-tracker was calibrated using a five-point calibration routine to obtain valid gaze positions in a stimulus-related coordinate system. The participant was lead to believe that exactly the same procedure was applied for the interactor. Subsequently, the real participant engaged in three interaction blocks to be familiarized with the task. After this practice session, remaining questions of the participant were answered. Both the participant and the interaction partner were then instructed not to communicate verbally with each other during the experiment and were asked to wear headphones in order "to prevent acoustical interferences" with their task performance. The eye-trackers were then recalibrated and the experiment started. After 28 of the 56 interaction blocks there was a 30 second break. Upon completion of the experiment, the partner was brought to another room while the participant was asked to fill out a brief questionnaire in which they had to indicate how difficult they had found the task on a 4-point scale, whether they had based their decision on considerations of human behavior or computers, whether they had used a certain strategy in the interaction, and whether there were specific criteria on which they based their decision. They were also asked to explicitly describe potential strategies and criteria. After completion of this questionnaire, all participants were debriefed and informed about the goal and purpose of the experiment. In total, the complete experimental session lasted approximately 50 minutes.

Data Analysis and Presentation

All data were analyzed using PASW Statistics 18 (SPSS Inc, Chicago, IL, www.spss.com). One-way ANOVAs for repeated measures were used to analyze the effect of the degree of gaze-following which was included in the analysis as a factor with seven levels. In order to be able to apply parametric statistics on proportional data, such as obtained from participant's judgments, all data were arcsine transformed [20,21]. Planned polynomial contrasts were applied for trend analysis. In addition to the main manipulation of the task, i.e. the systematic variation of the virtual character's gaze-following behavior, the gaze behavior of the participants themselves was analyzed to detect possible influences on the ascription of humanness. Whenever appropriate, i.e. for main effects and planned contrasts, omega squared (ω^2) is reported as a measure of effect size [22]. The following conventions for interpreting ω^2 are suggested: Small effects: $\omega^2<0.06$; Moderate effects: $\omega^2>0.06$ and $\omega^2<0.15$; Large effects: $\omega^2>0.15$ [23]. In the graphs representing the data, non-transformed data are used with error bars indicating the 95% confidence intervals. Post-experiment debriefing questionnaires were analyzed by an independent rater blind to the conditions of the study.

Results

Gaze Behavior of Participants

Before assessing the ascription of humanness based on the gaze reactions of the virtual character, we aimed at excluding potential effects of participants' own gaze behavior on performance. Two aspects of participants' gaze behavior were evaluated. In a first step, we investigated whether participants looked equally often to the left and right objects across conditions. This was clearly the case as indicated by left/right-ratios (Exp.1: 1.08, Exp.2: 1.04, Exp.3: 1.1, Exp.4: 1.04, Exp.5: 1.08) and supported by a one-way ANOVA comparing these ratios across conditions that did not yield any significant differences, $F(4, 108) = 2.08$, $p = .10$. In addition, the consistency of participants' gaze behavior was also taken into account. This is important because it is conceivable that whereas some participants alternate randomly between the left and right objects throughout interaction blocks, others chose to always fixate one of the two objects, thereby expressing higher consistency in their behavior. To assess the possibility that differences in consistency influence how participants experience the virtual character's gaze reactions and thus possibly their humanness rating, the longest chain of consecutive gaze shifts to the same object was extracted from each interaction block and used to calculate an average consistency index for each participant and each condition. An ANOVA comparing the average consistency across experiments did not yield any significant differences, $F(12, 408) = 1.11$, $p = 0.35$. Subsequently, the humanness ratings of each condition with the consistency index of that condition were correlated. The Pearson correlation coefficients were then included as a covariate in the repeated-measures ANOVAS employed for the within-group analyses of the effect of the independent variable (i.e. character's gaze-following behavior) on the dependent variable (i.e. the ascription of humanness) that will be described in the following sections.

Experiment 1: Interaction with a Naïve Confederate

In what we consider as the baseline task, the confederate was introduced as naïve to the participants' task. This means that he did not know that the real participant had to answer the question whether he had the impression to be interacting with another human or a computer program. We explicitly instructed participants that the confederate was unaware of the computer program randomly taking control of the virtual character's eye movements and of their task and thus could not knowingly help them in answering the question.

Participants. 26 healthy volunteers participated in this study ($M = 26.34$, $SD = 5.12$; 14 female). One female and one male participant needed to be excluded from the analysis due to technical problems during the experiment.

Results. The effects of increasing degrees of gaze-following on humanness ascription are depicted in Figure 3a. The results indicate that the proportion of human ratings increases with an increasing degree of gaze-following by the virtual character. A one-way repeated measures ANOVA including the degree of gaze-following as a factor with seven levels was performed on the data. Mauchly's test indicated that the assumption of sphericity was violated ($\chi^2 = 59.83$, $p<.001$). Degrees of freedom were therefore corrected by using the Greenhouse-Geisser estimates of sphericity ($\varepsilon = .54$). The results show a main effect of gaze-following on the ascription of humanness, $F(3.23, 74.34) = 5.31$, $p = .002$, $\omega^2 = 0.12$. Polynomial contrasts revealed a significant linear trend, $F(1, 24) = 13.54$, $p = .001$, $\omega^2 = 0.26$, thereby confirming the initial observation.

Discussion. Consistent with the literature on social gaze and social interaction, we hypothesized that participants would base their decision on congruent reactions to their own behavior. Indeed, the results show a highly significant linear trend and demonstrate that, when interacting with a putatively naïve confederate, participants' ratings in favor of a human interaction partner increased with increasing degrees of gaze-following. This indicates that during interaction with an unknown person there might be a default expectation of congruent reactions.

Experiment 2: Interaction with a Cooperative Interaction Partner

It has been argued that humans have a predisposition to interact cooperatively as soon as they interact [24,25]. To assess whether

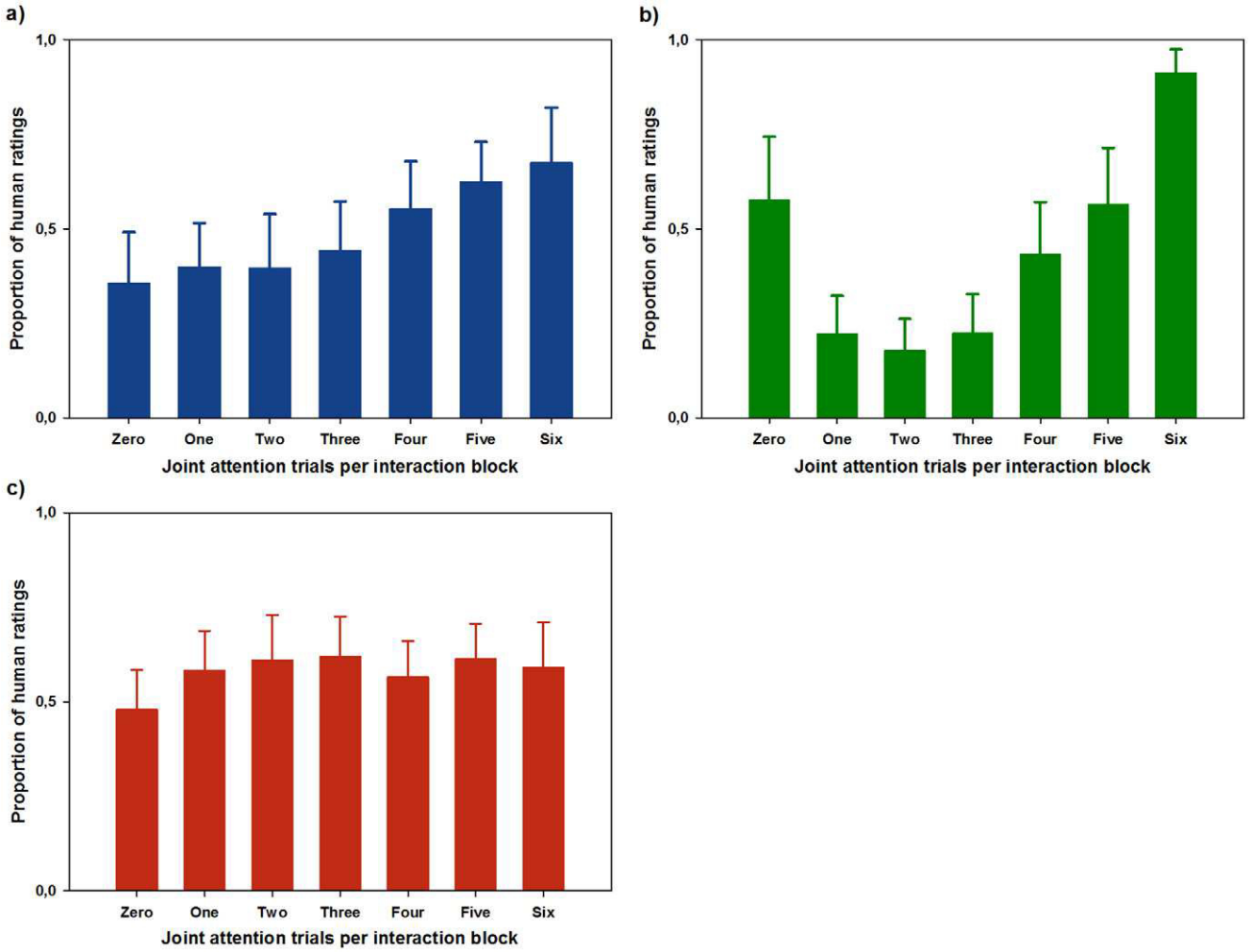

Figure 3. Experiments 1, 2, and 3: The ascription of humanness to a virtual character during interaction with an interactor that is *(a)* supposedly naïve to the participants' task, *(b)* introduced as cooperative, *(c)* or as competitive.
doi:10.1371/journal.pone.0027591.g003

the introduction of an explicitly cooperative context would either reinforce the congruency-based pattern of humanness ascription found in the previous experiment or would rather lead to a contingency-based pattern, we introduced the interaction partner as being aware of the participants' task in experiment 2. In addition, he was described as having been instructed to "cooperate", thus making the task as easy for the participant as possible. To stimulate a cooperative mindset, we also informed the participant that they both would receive additionally money if cooperation would lead to more correct decisions between human interactor and computer program.

Participants. 28 volunteers participated in this experiment (M = 26.96, SD = 6.65; 13 female). Two male participants were excluded because they did not believe the cover story.

Results. Figure 3b illustrates the mean responses for participants interacting with an interactor previously introduced as cooperative. Mean responses provide a first hint that during cooperative interaction the mere contingency seems to play an important role in humanness ascription. Again, Mauchly's test showed that the assumption of sphericity was violated ($\chi^2 = 80.92$, $p<.001$), and the Greenhouse-Geisser correction was used ($\varepsilon = .40$). Here, too, the degree of gaze-following had a highly significant effect on the ascription of humanness, $F(2.37, 59.3) = 22.63$; $p<.001$, $\omega^2 = 0.38$. There were highly significant linear, $F(1, 25) = 20.48$; $p<.001$, $\omega^2 = 0.20$, quadratic, $F(1, 25) = 38.3$; $p<.001$, $\omega^2 = 0.47$, and cubic, $F(1, 25) = 9.2$; $p = .005$, $\omega^2 = 0.05$, trends describing the u-shaped response pattern. A repeated-measures ANOVA including cooperativeness (experiment 1 vs. experiment 2) as a between-subjects factor showed that there was a significant difference in humanness ascription between experiments 1 and 2, $F(3.13, 150.17) = 7.04$; $p<.001$.

Discussion. Introducing the putative interaction partner as cooperative had a striking influence on the pattern of the ascription of humanness to the virtual character, which primarily followed a contingency-based pattern. Participants appear to discount the expectation of congruency of an interactor's reaction if the interactor is introduced as cooperative, indicating that in a cooperative context coordinated reactions seem to be more indicative of a human interactor than simple congruent reciprocation.

Experiment 3: Interaction with a Competitive Interaction Partner

This experiment assessed whether one of the prevalent response patterns from experiments 1 and 2 would still appear in a competitive situation. To this end, participants were informed that the confederate was aware of their task and instructed that he

should behave in a competitive way, hence making the decision as difficult as possible. To accentuate this manipulation, participants were told that they could earn extra amounts of money depending on their success rate. Conversely, the reimbursement of the other person was said to depend on his ability to trick the participant. It was hypothesized that participants would avoid the ascription of humanness in situations of maximal congruency or contingency of gaze reactions.

Participants. 21 healthy volunteers participated in this experiment (M = 29.9, SD = 4.95; 9 female).

Results. In Figure 3c the ascription of humanness in the presence of a competitive interactor is depicted. It is obvious that none of the previously described response patterns can be observed. Again, the assumption of sphericity was violated ($\chi^2 = 35.35$, $p = .02$) and the Greenhouse-Geisser correction ($\varepsilon = .57$) was employed in an ANOVA which did not show any significant effect of the degree of gaze-following on humanness ascription, $F(3.4, 61.16) = 1.11$; $p = .364$, and hence confirms the initial observation. Repeated measures ANOVAs including experiment (experiment 3 vs. experiment 1; experiment 3 vs. experiment 2) as a between-subjects factor demonstrated that humanness ascription during competitive interaction differed significantly from cooperative interaction, $F(3.31, 134.75) = 14.17$; $p<.001$, and showed a strong trend towards significance compared to the interaction with a naïve interactor, $F(3.53, 148.21) = 2.34$; $p = .056$.

Discussion. As predicted, when interacting with a competitive interactor, neither congruency nor mere contingency of reactions played a role in influencing the ascription of humanness. This demonstrates that participants expect a competitive partner to avoid reciprocation and coordination, thus further corroborating the importance of congruency and contingency in experiencing an interaction as an interaction with a human interactor.

Debriefing Questionnaires

For a better understanding of how participants addressed the task their responses in the post-experiment debriefing questionnaires were analyzed (see Figure S1). These questionnaires included four questions:

(1) Did participants base the ascription of humanness on considerations of human behavior or the function of a computer? Overall, the vast majority of participants based their ratings on considerations about human behavior (90.52%) rather than solely the function of computers (9.48%) while performing the task. This suggests that the non-verbal Turing test did not assess participant's hypotheses about how computers are programmed but indeed the experience of interaction with other persons.

(2) How difficult did participants rate the task on a scale from 1 (easy) to 4 (difficult)? The condition to which participants were assigned had a significant effect on their difficulty ratings, $F(2, 64) = 6.04$, $p = .004$. Tukey post-hoc comparisons of the three experiments revealed that difficulty ratings of participants who had interacted with a putatively cooperative interactor (i.e. Experiment 2) were significantly lower (M = 2.59, 95% CI [2.32, 2.88]) compared to difficulty ratings in the naïve (M = 3.1, 95% CI [2.86, 3.35]), $p = .017$, or competitive (M = 3.23, 95% CI [2.94, 3.51]), $p = .004$, condition. This indicates that the ascription of humanness was easiest for participants who had interacted with a cooperative interactor.

(3) Did participants use any behavioral strategy to unravel the nature of their interactor? An analysis of the presence of a strategy did not reveal any significant difference between the three conditions, $F(2, 67) = 1.84$, $p = .17$, indicating that the nature of the interaction partner did not have any effect on how strategic participants addressed the Turing test.

(4) Could participants report any specific criterion for deciding between having interacted between a human and a computer? A one-way ANOVA revealed that the condition had a significant effect on whether participants had a specific criterion for humanness ascription, $F(2, 67) = 10.99$, $p<.001$. Tukey post-hoc comparisons showed that participants who had interacted with a putatively competitive interactor had significantly fewer explicit criteria for humanness ascription (M = 0.10, 95% CI [0.01, 0.24]) compared to the naïve (M = 0.46, 95% CI [0.37, 0.78]), $p = .032$, or cooperative (M = 0.58, 95% CI [0.37, 0.78]), $p = .003$, condition. The proportion of explicit criteria did not differ between the naïve and the cooperative condition.

Eventually, we also looked at the comments in the questionnaires in a descriptive way. Notably, a considerable number of participants indicated that they expected a human interactor to either always follow their gaze or always avert their gaze and simply counted the occurrence of the expected reactions. In the following section two experiments including a concurrent cognitive load task will address the issue whether the Turing test provides a measure of strategic reasoning about humanness or rather of the implicit experience of an interaction as social.

Experiment 4: Interaction with a Naïve Confederate under Increased Cognitive Load

The possibility that participants simply test ad hoc hypotheses about human behavior in order to solve the Turing test provides a potential problem to our approach which aims at unraveling the factors that lead to the phenomenological experience of an interaction as an interaction with another human rather than strategic behaviors that might inform a decision between mind and machine. Social cognition has been distinguished from other domains of cognition by a high degree of automaticity and reflexivity of its core processes [26,27]. An increase of cognitive load in a so-called dual-task design is known to burden effortful *reflective* rather than automatic *reflexive* processes due to competition for limited cognitive resources [28]. In experiment 4 participants were instructed in the same way as in experiment 1. However, when the object changed color, a random number between 2 and 9 appeared superimposed on it. The concurrent cognitive load task consisted in adding up all six numbers that appeared during one interaction segment and to enter the sum after giving the response with respect to the nature of the interaction partner. We expected this manipulation to distract participants from any explicit strategy they could employ to inform the ascription of humanness.

Participants. 26 volunteers participated in this experiment (M = 25.85, SD = 3.3; 14 female). One participant needed to be excluded from the analysis because he did not believe the cover story.

Results. The results of humanness ascription during interaction with a naïve partner under cognitive load are depicted in Figure 4a. As in experiment 1, human ratings increase with increasing gaze-following. Mauchly's test indicated that the assumption of sphericity was violated ($\chi^2 = 90.23$; $p<.001$) and degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon = .36$). The results of a one-way repeated-measures ANOVA indicated a highly significant effect of gaze-following on humanness ascription, $F(2.16, 47.51) = 10.45$, $p<.001$, $\omega^2 = 0.24$. Polynomial contrasts revealed both a highly significant linear, $F(1, 20) = 12.87$, $p = .001$, $\omega^2 = 0.29$, and quadratic trend, $F(1, 20) = 11.09$, $p = .001$, $\omega^2 = 0.16$, as in experiment 1. A repeated measures ANOVA including experiment as a between-subjects factor (experiment 4 vs. experiment 1) showed that humanness ascription during interaction with a naïve partner was not significantly affected by the presence of a

concurrent cognitive load task, $F(2.89, 137.51) = 0.59$, $p = .62$, and thus confirmed the results from experiment 1.

Participants of experiment 4 were eventually separated into a high- and a low-performance group by means of a median split based on cognitive load performance. A one-way repeated measures ANOVA with performance group as a between-subjects factor demonstrated a significant effect of cognitive load performance on the ascription of humanness, $F(2.71, 56.86) = 2.88$, $p = .049$. Polynomial trend analysis within these two groups indicates that the high-performance group shows a stronger linear trend, $F(1, 11) = 21.9$, $p = .001$, $\omega^2 = 0.59$, compared to the low-performance group, $F(1, 11) = 6.84$, $p = .024$, $\omega^2 = 0.28$. This is illustrated by figure 4a which demonstrates that participants in the high-performance group (solid lines) show a much more pronounced congruency-based response pattern than those in the low-performance group.

Discussion. Overall, humanness ascription in the naïve condition did not change significantly under concurrent cognitive load. However, splitting participants into a low- and a high-performance group demonstrated a clear effect of the load manipulation: Participants who obtained higher scores in the load task showed a more pronounced linear trend in humanness ascription, that is, they based their ratings maximally on the congruency of the virtual character's reaction. As higher performance in the cognitive load task is indicative of greater distraction by this task, these results emphasize that the congruency of gaze-reactions is the most prominent cue for humanness ascription when cognitive resources are burdened during the Turing test. This can be taken to suggest that congruency-based responses are produced in an implicit and automatic fashion rather than being a product of strategic reasoning processes.

Experiment 5: Interaction with a Cooperative Interaction Partner under Increased Cognitive Load

This experiment followed the same rationale as experiment 4 and assessed the effect of concurrent cognitive load on humanness ascription during interaction with a cooperative interactor.

Participants. In this experiment, 29 healthy volunteers participated (M = 25.11, SD = 4.42; 17 female). One male and a female participant were excluded from the analysis due to technical problems during the experiment.

Results. *Effect of Gaze Reactions.* As in experiment 3, the mean responses suggest that again overall contingency seems to play an important role in the experience of an interaction as social (Figure 4b). The Greenhouse-Geisser correction ($\varepsilon = .47$) was used to correct for the violation of sphericity as indicated by Mauchly's test ($\chi^2 = 84.24$, $p < .001$). Again, the degree of gaze-following had a highly significant effect on humanness ascription, $F(2.79, 72.64) = 12.52$, $p < .001$, $\omega^2 = 0.29$, and displayed significant linear, $F(1, 26) = 6.03$; $p < .021$, $\omega^2 = 0.05$, and quadratic trends, $F(1, 26) = 25.42$, $p < .001$, $\omega^2 = 0.44$. As indicated by a repeated measures ANOVA including the presence of the cognitive load task as a between-subjects factor (experiment 5 vs. experiment 2) the addition of a concurrent cognitive load task did not lead to group differences in humanness ascription, $F(2.75, 134.57) = 1.22$, $p = .31$.

Participants again were separated into high- and low-performers by a median split of cognitive load performance. Unlike experiment 4 including load performance as a between-subjects factor did not yield any significant effect, $F(2.65, 63.68) = 0.36$, $p = .36$. In contrast, humanness ascription differed significantly between the two cognitive load experiments (experiment 4 vs. experiment 5), $F(3.12, 146,47) = 3.81$, $p = .011$, thus indicating that the difference in response patterns observed in naïve compared to cooperative interactions remained consistent despite the addition of a cognitive load task.

Discussion. The results of this experiment confirmed that humanness is ascribed based on the mere contingency of gaze reactions when the Turing test is performed with a cooperative interactor. Both high- and low-performers equally ascribed humanness based on contingent rather than congruent responses, indicating that contingency is the prevalent cue irrespective of the degree of cognitive burdening imposed by the cognitive load task. The cooperative interaction hence seems to

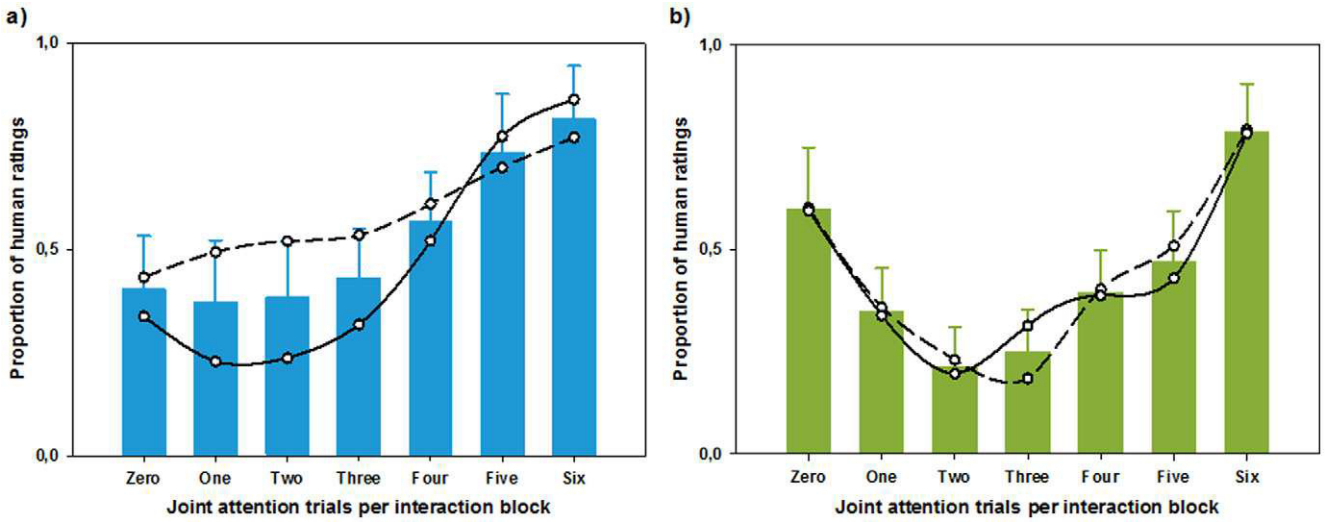

Figure 4. Experiments 4 and 5: The ascription of humanness to a virtual character while concurrently solving a cognitive load task. A median split separated participants with high and low scores in the cognitive load task. Solid lines represent the mean humanness ratings of high performers, whereas dashed lines represent low performers. *(a)* During naïve interaction cognitive load performance had an effect on humanness ascription ($p = 0.49$). High performers show a stronger congruency-based response pattern compared to low performers. *(b)* In cooperative interaction there was no effect of load performance on the ascription of humanness.
doi:10.1371/journal.pone.0027591.g004

induce an implicit expectation of contingency that is not altered by any strategic reasoning.

Further Hints to the Implicitness of Humanness Ascription

In the two cognitive load experiments the focus of the manipulation was during the interaction phase. The rationale was that the task would distract people from thinking about the interaction process and engaging in strategic reasoning about the task. Nevertheless, the decision between human or computer might not emerge *during* but completely *after* the interaction. To address this possibility we analyzed reaction times (see Figure 5d). A one-way ANOVA including all experiments was performed and demonstrated a main effect of experimental group on reaction times, $F(4, 116) = 3{,}79$, $p = .006$. Pooling the data into load and no-load experiments showed that this effect was due to significantly higher reaction times in the load ($M = 2250.7$, $SE = 94.4$) compared to the no-load ($M = 1877.81$, $SE = 70.23$) tasks, $t(119) = -2{,}56$, $p = .012$, suggesting higher cognitive demands resulting from the combination of the humanness ascription and the cognitive load task. A one-way ANOVA did not reveal any significant differences of reaction times between the no-load conditions (experiments 1, 2, and 3), $F(2, 68) = 2.01$, $p = .142$. A comparison of the two load experiments (experiments 4 and 5) also did not show any significant difference, $t(48) = .92$, $p = .364$. Although this suggests that the decision is made during the interaction, it cannot be ruled out that reasoning processes between the end of the interaction block and the button press play a role in humanness ascription.

To investigate this matter, a median split of reaction times was performed for all experiments (Figure 5a–e). In the naïve condition, a one-way repeated-measures ANOVA revealed a significant effect of reaction time on humanness ascription, $F(3.09, 70.99) = 3.02$, $p = .034$, $\omega^2 = 0.12$. Separate ANOVAs for participants with short and long reaction times showed that the degree of gaze-following only had an effect in the fast responders, $F(2.82, 33.79) = 5.39$, $p = .004$, $\omega^2 = 0.32$, who showed a highly significant linear trend of humanness ascription, $F(1, 12) = 13.02$, $p = .004$, $\omega^2 = 0.32$. In the slow responders, there was no such effect, $F(3.4, 37.36) = 1.23$, $p = .31$. In the naïve condition including cognitive load an ANOVA revealed an effect of reaction time on the ascription of humanness, $F(2.52, 55.48) = 5.44$, $p = .004$, $\omega^2 = 0.22$. Similar results as in the naïve condition without cognitive load were indicated by separate ANOVAs for slow and fast responders. Gaze-following only had a significant effect in the fast responders, $F(1.74, 20.83) = 17.13$, $p < .001$, $\omega^2 = 0.44$, where also a linear trend was present, $F(1, 12) = 24.04$, $p < .001$, $\omega^2 = 0.43$, but not in the slow responders, $F(3.36, 33.55) = 1.8$, $p = .11$. This suggests that the longer participants think about their decision after the interaction, the lesser they take congruency into account as a humanness cue. Interestingly, there was no such an effect for experiments 2, 3, and 5, indicating that during cooperative interaction, the ascription of humanness is implicitly based on the contingency of gaze reactions without being influenced by the time spent on thinking about the decision.

The reaction time data are supported by participant's responses to the questions whether they had behavioral strategies and whether they could mention explicit criteria for humanness ascription. Concerning the question whether they had used specific strategies to investigate whether they had interacted with another human or a computer (Figure S1c), this was significantly less the case in the experiments including a concurrent cognitive load task, $\chi^2(1) = 6.23$, $p = .013$. In addition, although this was only a statistical trend, participants did report specific criteria for the ascription of humanness (Figure S1d) considerably less often in the cognitive load experiments compared to the experiments with increased cognitive load, $\chi^2(1) = 3.27$, $p = .07$. These results indicate that manipulation of cognitive load was successful in reducing strategic behavior of participants as well as their awareness of specific criteria for the ascription of humanness.

Discussion

In a series of experiments, we have made use of a novel interactive eye-tracking paradigm to establish what we describe as a non-verbal Turing test. This setup makes it possible to assess parameters of gaze-based interaction which lead to the experience of a truly social encounter with a real human interaction partner. Hereby we could overcome the paradoxical situation of previous studies on social cognition in which the behavior of a single person is observed in isolation from others. The experience of being involved in interaction is constituted by two aspects: Firstly, participants in our experiments experience that they are directly addressed by the virtual character whose gaze behavior is made contingent on their own in real time. The necessity of "being addressed as you" has recently been advanced as a second-person approach to social cognition in the fields of social cognition and neuroscience [5,29,30]. Secondly, the paradigm enables participants to directly observe the consequences of their actions on another agent as it would occur in real-life interaction. This is vital for making sense of one's own behavior in an interactive context and for its adjustment to situational requirements.

This newly developed approach provides important and novel insights on the process underlying the ascription of humanness to virtual characters in social encounters. In order to model different social contexts, participants engaged in the non-verbal Turing test under changing situational demands: Experiment 1 assessed humanness ascription during interaction with an interactor who was thought to be *naïve* to the task in order to assess the default ascription pattern when there is no knowledge about the interactor. Consistent with our hypothesis, participants activated a congruency-based expectation and increasingly ascribed humanness to the virtual character with increasing degrees of gaze-following. The results of experiment 2 demonstrate that this pattern can be modulated depending on the previous knowledge about the behavioral predisposition of the interaction partner and changes to a contingency-based analysis of behavior in the presence of a *cooperative* partner. As predicted, experiment 3 showed that the ascription of humanness during interaction with a *competitive* interactor was neither based on congruency nor on contingency of gaze reactions.

The Special Case of Gaze

Before turning to an in-depth discussion of our results there are two controversial issues related to the operationalization of the interaction process using gaze cues and to the resulting explanations that need to be addressed.

First of all, it might be argued that gaze-following is merely a form of motor mimicry which refers to a subtle imitation of the behavior of an interaction partner. Consequently, the ascription of humanness might rely on mimicry-related processes which are known to increase rapport, empathy, and liking between mimicker and mimickee and thereby result in increased bonding of the interactors [31]. Although gaze-following naturally has an imitative component, motor mimicry can clearly be distinguished from gaze-following in a number of respects. Chartrand and Bargh [32] describe mimicry as non-conscious imitation "such that one's behavior passively and unintentionally changes to match that of

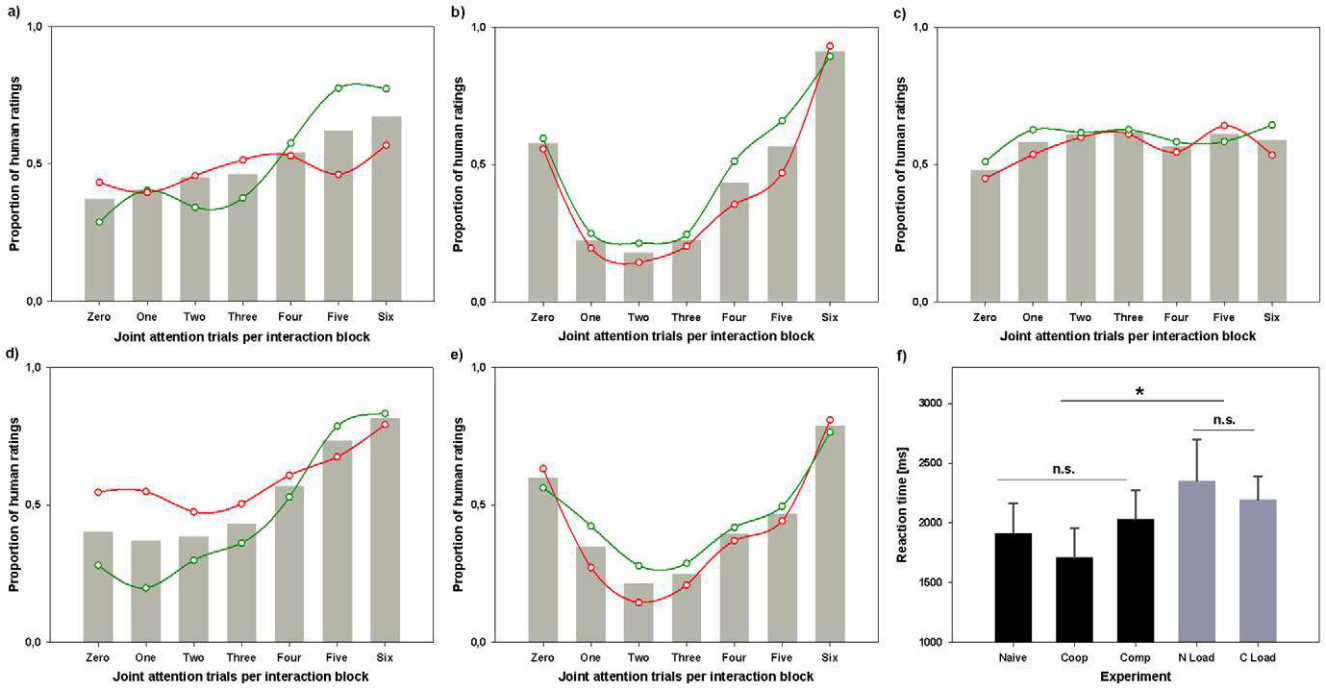

Figure 5. Reaction times of humanness decisions split by median. Grey bars indicate mean ratings. Mean ratings of fast responders (reaction time below median) are indicated by green scatter plot, mean ratings of slow responders are indicated by red scatter plot. Effects of response time are indicated in brackets. *(a)* Naïve interactor ($p = .034$): In fast responders humanness ascription is driven more strongly by congruency than in slow responders. *(b)* Cooperative interactor (n.s.). *(c)* Competitive interactor (n.s.). *(d)* Naïve interactor + cognitive load ($p = .004$): Fast responders show stronger congruency-based response patterns compared to slow responders. *(e)* Cooperative interactor + cognitive Load (n.s.). *(f)* Mean reaction times for all experiments (see text for details).
doi:10.1371/journal.pone.0027591.g005

others in one's current social environment'' (p.893). The involvement of a distinct task as in our series of experiments makes it difficult to argue that the other's gaze-following is passive or unintentional. Another important argument is that the appraisal of the mimicker decreases and other positive social effects break down once the mimickee becomes aware of being mimicked, possibly because the imitative behavior is evaluated as an intentional expression of conformity directed at the attainment of reward or approval [33,34]. In our task it is obvious that the participant is aware of the other following or not following his gaze, as this is the criterion on which the decision between human and computer is based. A further distinction concerns the *function* of mimicry and gaze-following. Whereas the main function of mimicry seems to be the general facilitation of dyadic interactions, gaze-following is related to triadic rather than dyadic interactions where it serves the purpose of keeping track of another person's focus of attention, thereby paving the way to an understanding of this person's mental states [9,10]. The distinct task structure, participants' awareness of the other's reactions, as well as the functional role of gaze-following clearly argue against any substantial role of mimicry in the present study.

Secondly, a critical reader might ask whether the effects demonstrated here are gaze-specific or whether they could potentially be replicated using different channels of non-verbal behavior. Undoubtedly, some of the effects reported in this article might appear in a Turing-test-like study involving other forms of interaction. However, the aim of the present study was to uncover the basic aspects of 'online' social interaction that lead to the experience of this interaction as an interaction with another human. In order to obtain a valid operationalization, we identified and aimed at fulfilling two main criteria without which the task could not provide a valid experimental investigation of online social interaction. First, the task needs to provide a high level of ecological validity, i.e. both channel and process of the interaction must be highly salient in everyday social interactions. Second, the task must provide a high degree of experimental control. Obviously, other cue system could be used to model contingency and congruency of an interaction in an experimentally controllable fashion. For example, a similar study design could involve pressing a button, moving a cursor, producing a sound or any combination of these cues. However, this would not satisfy the criterion of ecological validity as these activities are not part of every-day social interactions. Furthermore, social gaze is distinct from other communicative channels in one crucial aspect. Already more than 40 years ago, Gibson and Pick noted that gaze ''can be treated as a source of stimulation as well as a type of response. The eyes not only look but are looked at'' ([13], p.386) and that hence in the act of looking perception and action are inseparable. Taken together, for the following reasons, social gaze seems most ideally suited for a Turing-test-like assessment of social interaction: (i) It readily occurs in natural interaction, (ii) it is linked to an understanding of other's minds, (iii) it is easily controllable in an experimental setting, and (iv) it combines stimulation and response in one action.

The Valence of Gaze Aversion and Gaze-Following

As a key finding, our studies demonstrate that human beings who interacted with a putatively naïve partner displayed an implicit expectation of gaze-following behavior and experienced an interaction as social when the interactor followed their gaze and engaged in joint attention with them.. This effect is surprisingly robust given that the only piece of information available to the

participants was that the partner had been instructed to react to their gaze by "freely choosing to look to the same or the other object" without being able to willingly help them to solve the task.

Why is maximally averted gaze not indicative of a human interactor? In the first instance, this might be related to the fact that in the domain of social gaze valence is an inherent property of the contingency continuum which ranges from maximal gaze aversion to maximal gaze-following [6]. The neglect of maximal gaze aversion as a cue to humanness during interaction with a naïve interactor might be related to the negative valence of gaze aversion that has been demonstrated on various levels. For example, in a study on the effects of gaze cues on person construal it has been shown that participants produced higher ratings of both likeability and attractiveness for pictures of people shifting the gaze towards them compared to pictures of people averting their gaze from them [35]. In another study [36], participants viewed video sequences displaying a human face either directing its gaze at them or averting it by looking left or right from time to time. As a between-subjects factor the degree of gaze aversion was varied. After having watched the movies, participants had to fill out a social rejection questionnaire which showed that feelings of exclusion and ostracism increased with increasing total duration of gaze aversion. In addition, gaze aversion generally increased feelings of negative mood and decreased prosocial attitudes. Additional evidence for the negative valence of gaze aversion comes from an EEG experiment in which participants viewed live faces displaying either direct or averted gaze [36]. An analysis of EEG activity revealed that direct gaze elicited left-hemispheric frontal activation which has been related to approach motivation. On the contrary, averted gaze resulted in right-sided frontal activation that has been related to an avoidance motivation, suggesting that gaze aversion triggers neural responses related to negative affect [37].

Is there comparable support for a positive valence of gaze-following and joint attention? A crucial distinction has been made between other- and self-initiated joint attention. One can either respond to bids of joint attention by others or initiate joint attention by leading someone's gaze. Whereas gaze-following has been observed in other species, the ability and spontaneous motivation to lead someone's gaze is uniquely human. Its function is to share interests and pleasant experiences regarding objects in the environment with others [10]. For the present study, a recently discovered motivational aspect of self-initiated joint attention is of great importance. Schilbach and colleagues [16] report that being involved in joint attention, irrespective of its initiator, results in the activation of regions of the so-called "social brain", such as the medial prefrontal cortex. This region has been implicated in mentalizing, i.e. in thinking about other person's goals and intentions [38]. Initiating joint attention oneself, however, is associated with increased neural activity in the ventral striatum as part of the brain's reward system whose activity changes have been linked to hedonic experiences and the anticipation of reward [39,40]. In addition, there was a significant correlation of the strength of striatal activation with ratings of the pleasantness of joint attention obtained in a post-scan questionnaire. These findings indicate that self-initiated joint attention triggers reward-related processing and hence provides an intrinsic motivation for engaging others in joint attention. In other words, we seek for reciprocation and enjoy being able to elicit congruent responses from others to our actions. Taken together, we believe that these positive connotations of gaze-following may be crucial in informing the ascription of humanness.

From Joint Attention to Joint Action by Cooperation

Our results provide compelling evidence for the significant impact of prior knowledge about the goal of the presumed interactor on the experience of an interaction. When the interactor was explicitly introduced as cooperative, the ascription of humanness was not based on congruency but rather followed the actual contingency of the virtual character's reactions more closely. This finding is consistent with our hypothesis and indicates that people were, in fact, not blind to the actual contingencies, but only integrate them when the interactor's disposition to cooperate is known.

How can cooperation lead to the discounting of the expectation of congruent gaze reactions? Cooperation, in the traditional view, is a behavior that is selected to provide mutual benefit to both the actor and the recipient. Cooperation often requires that immediate benefits are discounted in order to gain a delayed reward [41,42]. However, cooperation has not only been defined in terms of its fitness consequences, but also in a mechanistic sense as a form of behavioral coordination [18]. In this definition, particular emphasis is put on the necessity of coordination between the cooperative interaction partners which is regarded as an "important proximate mechanism needed to accomplish cooperation" ([43], p. 7). Interestingly, the coordination of behaviors is not only pivotal for cooperation, but also for joint action [44]. For example, musicians playing instruments in a band, a couple dancing together, or construction workers building a house demonstrate cases of joint action. It is hence possible that the discounting of mere congruency in the cooperative condition is a consequence of participants interpreting the interaction as a form of joint action. An analysis of the degree of coordination expected by participants from a human interactor and an assessment of the criteria that an interaction needs to fulfill in order to be classified as a joint action might help to assess this option.

Does cooperative interaction in the non-verbal Turing test qualify as a joint action? There are two salient coordinated behavioral patterns that occur in the Turing test, namely maximal gaze-following or maximal gaze aversion. Data from the naïve condition suggest that maximal gaze-following constitutes the most basic and effortless form of coordinative behavior which seems to be expected "by default" when people engage in interaction. In the cooperative situation, any form of contingency is judged as indicative of a human interaction partner, thus indicating that participants expect a higher degree of coordination. This strong expectation of coordinated behavior irrespective of the congruency of reactions might be taken to suggest that participants understand the cooperative interaction as a situation of joint action. Fiebich and Gallagher [45] have recently identified three conditions that need to be satisfied before interactors can be said to be engaged in joint action: i) they need to have a shared goal or intention, ii) they must have common knowledge of aiming at this goal together, and iii) they have to participate in coordinated patterns in order to reach this goal. These criteria are fulfilled in the cooperative version of the Turing test: (i) The shared goal of increasing the common monetary reward is easily identified for the interaction with a cooperative interaction partner. (ii) As this has been communicated explicitly, the participant can also assume that they are aiming at this goal together. (iii) The contingency-driven response pattern indicates that participants strongly expected the other to coordinate his behavior to their actions on a higher level than mere congruency. We speculate that this demonstrates an intrinsic expectation of higher-order coordination in cooperation compared to the unrestrained interaction format in the naïve condition and thus provides evidence that the interaction with a cooperative interactor is automatically interpreted as a situation of joint action.

Experiencing Interaction or Thinking about Interaction?. It might be argued that the ascription of humanness could have been based on reasoning processes which are not

related to the experience of social interaction. Social cognition has been described as being largely constituted by automatic processes are fast, unconscious, and do not require willful regulatory efforts [27,46]. Hence, if participant's judgments were the outcome of conscious, deliberate, and strategic thought processes this would pose a problem to our claim of presenting these judgments as measures of the experience of interacting with another human. We assessed this possibility in several respects. First of all, the addition of a concurrent cognitive load task in experiments 4 and 5 specifically aimed at interfering with strategic processes during the interaction process by burdening the cognitive system of participants.

The results of these experiments clearly demonstrated that during naïve interaction the increase of cognitive load lead to an increased in congruency-based humanness ascription. Notably, participants who obtained high scores in the cognitive load task based the ascription of humanness more strongly on congruency than participants with low scores. This indicates that the interference created by the load task unraveled implicit or automatic response patterns. In cooperative interaction, on the other hand, the presence of the cognitive load task had no effect on humanness ascription, demonstrating that contingency-based responses represent implicit judgments of humanness. Overall, participants in the experiments including cognitive load reported that they used less strategies and less explicit criteria of humanness ascription, thereby further corroborating the effectiveness of the load manipulation. Considering that the decision between human and computer could take place completely after the interaction itself, reaction times were analyzed by splitting participants into fast and slow responders. In interactions with a naïve interactor, irrespective of the presence of a cognitive load task, fast responders base humanness ascription more strongly on congruency than slow responders. Taken together, these findings indicate that we were able to address the implicit processes leading to the experience of an interaction as an interaction with a human agent rather than results of careful deliberation that might inform a decision between mind and machine.

Outlook and Conclusions

Insights into how congruency and contingency of reactions to our own gaze behavior lead to the experience of an interaction as social address the interests of various fields of research. For instance, the current paradigm is likely to provide a useful tool to investigate impairments of the ability to engage in online social interaction in psychiatric disorders, such as it is observed in schizophrenia and autism [2]. The current methodological developments and empirical results could also inform research on human-computer interfaces aiming at the development of virtual agents that appear and behave human in a natural way in order to facilitate smooth interaction [47,48]. Clearly, such developments can benefit from research unraveling the core aspects of human social interaction by using truly interactive paradigms. Most obviously, however, the adaptation of the present experimental design for neuroimaging studies will provide a powerful tool for the study of the neural underpinnings of social interaction. In this respect, it can be hypothesized that gaze-based interaction with a naïve confederate might lead to an increase in neural activity in areas of the mentalizing system such as the medial prefrontal cortex [49]. In addition, conditions with highly congruent reactions might correlate with increased activity in brain areas implicated in the processing of reward such as the amygdala and the ventral striatum [16,50,51]. While competitive interaction might also concur with an increase of neural activity in mentalizing areas it would be interesting to investigate whether the competitive context could also lead to a decrease of activity in reward-related neurocircuitry when observing joint attention. Likewise, an interesting question concerns the neural substrates of contingency evaluation in a cooperative context: Does the presence of a shared goal lead to a decrease of activity in the mentalizing system in favor of activation of brain areas implicated in coordinated behavior (e.g., [52])? Furthermore, it will be interesting to investigate whether changes in activation of the reward system in response to positively contingent gaze-reactions could generalize to contingent reactions irrespective of their valence depending on the situational context.

In summary, our results demonstrate that the use of innovative methodology and experimental designs makes it possible to address the interaction process itself instead of focusing on the study of single minds in isolation [1]. Though still rare, truly interactive paradigms have also been advanced by other researchers in psychology and cognitive neuroscience [1,53–56]. This emphasizes the need for such studies if we want to understand why and how we interact with others in a more sophisticated way than any other species.

Supporting Information

Figure S1 Overview of participants' responses to the post-experiment debriefing questionnaire.
(TIF)

Acknowledgments

The authors would like to thank the members of the Neuroimaging Group of the Department of Psychiatry and Psychotherapy for providing helpful support and insightful comments on previous versions of this article.

Author Contributions

Conceived and designed the experiments: UJP KV GB. Performed the experiments: UJP. Analyzed the data: UJP BT LS. Contributed reagents/materials/analysis tools: LS UJP GB. Wrote the paper: UJP LS KV BT.

References

1. Becchio C, Sartori L, Castiello U (2010) Toward You: The Social Side of Actions. Current Directions in Psychological Science 19: 183–188.
2. Schilbach L, Eickhoff SB, Cieslik EC, Kuzmanovic B, Vogeley K (2011) Shall we do this together? Social gaze influences action control in a comparison group, but not in individuals with high-functioning autism. Autism.
3. Wheeler M (2005) Reconstructing the cognitive world: the next step. MIT Press.
4. De Jaegher H (2009) Social understanding through direct perception? Yes, by interacting. Conscious Cogn 18: 535–542.
5. Schilbach L (2010) A second-person approach to other minds. Nat Rev Neurosci 11: 449.
6. Argyle M, Cook M (1976) Gaze and mutual gaze. Cambridge University Press.
7. Senju A, Johnson MH (2009) The eye contact effect: mechanisms and development. Trends Cogn Sci (Regul. Ed.) 13: 127–134.
8. Emery NJ (2000) The eyes have it: the neuroethology, function and evolution of social gaze. Neurosci Biobehav Rev 24: 581–604.
9. Shepherd SV (2010) Following gaze: gaze-following behavior as a window into social cognition. Front Integr Neurosci 4: 5.
10. Mundy P, Newell L (2007) Attention, Joint Attention, and Social Cognition. Curr Dir Psychol Sci 16: 269–274.
11. Moll H, Tomasello M (2007) Cooperation and human cognition: the Vygotskian intelligence hypothesis. Philos Trans R Soc Lond., B, Biol Sci 362: 639–648.
12. Hardin CD, Higgins ET (1996) Shared Reality: How social verification makes the subjective objective. In: Higgins ET, Sorrentino RM, eds. Handbook of motivation and cognition: The interpersonal context. New York: Guilford, Vol. 3 p.
13. Gibson JJ, Pick AD (1963) Perception of another person's looking behavior. Am J Psychol 76: 386–394.

14. Wilms M, Schilbach L, Pfeiffer U, Bente G, Fink GR, et al. (2010) It's in your eyes—using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. Soc Cogn Affect Neurosci 5: 98–107.
15. Bente G, Eschenburg F, Krämer NC (2007) Virtual Gaze. A Pilot Study on the Effects of Computer Simulated Gaze in Avatar-Based Conversations. In: Shumaker R, ed. Virtual Reality. BerlinHeidelberg: Springer Berlin Heidelberg 4563: 185–194.
16. Schilbach L, Wilms M, Eickhoff SB, Romanzetti S, Tepest R, et al. (2010) Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. J Cogn Neurosci 22: 2702–2715.
17. Turing A (1950) Computing Machinery and Intelligence. Mind LIX: 433–460.
18. Taborsky M (2007) Cooperation built the Tower of Babel. Behav. Processes 76: 95–99.
19. Elson M, Bente G (2009) CORTEX - Computer-Aided Registration Tool for Experiments. Cologne, Germany, Available at: http://cortex.uni-koeln.de.
20. Snedecor GW, Cochran WG (1989) Statistical Methods. 8th ed. Iowa State University Press.
21. McDonald JH (2009) Handbook of Biological Statistics. 2nd ed. BaltimoreMD: Sparky House Publishing.
22. Streiner DL, Norman G (2008) Biostatistics: The Bare Essentials. 3rd ed. Hamilton, ON: BC Decker Inc.
23. Cohen J (1988) Statistical power analysis for the behavioral sciences. Routledge.
24. Tomasello M, Carpenter M, Call J, Behne T, Moll H (2005) Understanding and sharing intentions: the origins of cultural cognition. Behav Brain Sci 28: 675–691; discussion 691-735.
25. Tomasello M, University S (2009) Why We Cooperate. The MIT Press.
26. Lieberman MD (2007) Social Cognitive Neuroscience: A Review of Core Processes. Annu Rev Psychol 58: 259–289.
27. Bargh JA, Williams EL (2006) The Automaticity of Social Life. Curr Dir Psychol Sci 15: 1–4.
28. Baddeley A, Hitch G (1974) Working Memory. In: Bower GH, ed. The psychology of learning and motivation: Advances in research and theory. New York: Academic Press 8: 47–89.
29. Reddy V (2003) On being the object of attention: implications for self-other consciousness. Trends Cogn Sci (Regul. Ed.) 7: 397–402.
30. McGann M, De Jaegher H (2009) Self–other contingencies: Enacting social perception. Phenom Cogn Sci 8: 417–437.
31. Chartrand TL, van Baaren R (2009) Chapter 5 Human Mimicry. In: Academic Press, Vol. Volume 41. pp 219–274.
32. Chartrand TL, Bargh JA (1999) The chameleon effect: the perception-behavior link and social interaction. J Pers Soc Psychol 76: 893–910.
33. Thelen MH, Frautschi NM, Roberts MC, Kirkland KD, Dollinger SJ (1981) Being imitated, conformity, and social influence: An integrative review. Journal of Research in Personality 15: 403–426.
34. Lakin JL, Chartrand TL (2003) Using Nonconscious Behavioral Mimicry to Create Affiliation and Rapport. Psychological Science 14: 334–339.
35. Mason MF, Tatkow EP, Macrae CN (2005) The look of love: gaze shifts and person perception. Psychol Sci 16: 236–239.
36. Hietanen JK, Leppänen JM, Peltola MJ, Linna-aho K, Ruuhiala HJ (2008) Seeing direct and averted gaze activates the approach-avoidance motivational brain systems. Neuropsychologia 46: 2423–2430.
37. Harmon-Jones E, Gable PA, Peterson CK (2010) The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update. Biological Psychology 84: 451–462.
38. Amodio DM, Frith CD (2006) Meeting of minds: the medial frontal cortex and social cognition. Nat. Rev. Neurosci 7: 268–277.
39. Spicer J, Galvan A, Hare TA, Voss H, Glover G, et al. (2007) Sensitivity of the nucleus accumbens to violations in expectation of reward. Neuroimage 34: 455–461.
40. Liu X, Powell DK, Wang H, Gold BT, Corbly CR, et al. (2007) Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. J. Neurosci 27: 4587–4597.
41. Fehr E, Rockenbach B (2004) Human altruism: economic, neural, and evolutionary perspectives. Curr Opin Neurobiol 14: 784–790.
42. Brosnan SF, Salwiczek L, Bshary R (2010) The interplay of cognition and cooperation. Philos Trans R Soc Lond., B, Biol Sci 365: 2699–2710.
43. Noë R (2006) Cooperation experiments: coordination through communication versus acting apart together. Animal Behaviour 71: 1–18.
44. Knoblich G, Sebanz N (2008) Evolving intentions for social interaction: from entrainment to joint action. Philos Trans R Soc Lond., B, Biol Sci 363: 2021–2031.
45. Fiebich A, Gallagher S. (in revision) Joint Attention in Joint Action. Philosophical Psychology).
46. Bargh JA, Chartrand TL (1999) The unbearable automaticity of being. American Psychologist 54: 462–479.
47. Burgoon JK, Bonito JA, Bengtsson B, Cederberg C, Lundeberg M, et al. (2000) Interactivity in human-computer interaction: a study of credibility, understanding, and influence. Computers in Human Behavior 16: 553–574.
48. Epley N, Waytz A, Cacioppo JT (2007) On seeing human: a three-factor theory of anthropomorphism. Psychol Rev 114: 864–886.
49. Frith CD, Frith U (2006) The neural basis of mentalizing. Neuron 50: 531–534.
50. Adolphs R (2010) What does the amygdala contribute to social cognition? Ann N Y Acad Sci 1191: 42–61.
51. Krach S, Paulus FM, Bodden M, Kircher T (2010) The rewarding nature of social interactions. Front Behav Neurosci 4: 22.
52. Kokal I, Gazzola V, Keysers C (2009) Acting together in and beyond the mirror neuron system. Neuroimage 47: 2046–2056.
53. Fliessbach K, Weber B, Trautner P, Dohmen T, Sunde U, et al. (2007) Social comparison affects reward-related brain activity in the human ventral striatum. Science 318: 1305–1308.
54. Meulenbroek RGJ, Bosga J, Hulstijn M, Miedl S (2007) Joint-action coordination in transferring objects. Exp Brain Res 180: 333–343.
55. Redcay E, Dodell-Feder D, Pearrow MJ, Mavros PL, Kleiner M, et al. (2010) Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. Neuroimage 50: 1639–1647.
56. Richardson DC, Dale R, Kirkham NZ (2007) The art of conversation is coordination: common ground and the coupling of eye movements during dialogue. Psychol Sci 18: 407–413.

Title:

Why we interact: On the functional role of the ventral striatum during real-time social interactions

Authors:

Ulrich J. Pfeiffer[1]*, Leonhard Schilbach[1,2], Bert Timmermans[1], Bojana Kuzmanovic[3], Alexandra L. Georgescu[1], Gary Bente[4], Kai Vogeley[1,5]

Affiliations:

[1]Neuroimaging Lab, Department of Psychiatry and Psychotherapy, University Hospital Cologne, Kerpener Strasse 62, 50937 Cologne

[2]Max-Planck-Institute for Neurological Research, Gleueler Strasse 50, 50931 Cologne

[3]Institute of Neuroscience and Medicine – Ethics in the Neurosciences (INM8), Research Center Juelich, Leo-Brandt-Strasse, 52428 Juelich

[4]Institute of Social and Media Psychology, Department of Psychology, Faculty of Human Sciences, University of Cologne, Richard-Strauss-Strasse 2, 50931 Cologne

[5]Institute of Neuroscience and Medicine – Cognitive Neuroscience (INM3), Research Center Juelich, Leo-Brandt Strasse, 52428 Juelich

*Correspondence to: ulrich.pfeiffer@uk-koeln.de

Abstract:

There is ample evidence that human primates have a need for social contact and experience interactions with conspecifics as intrinsically rewarding. However, little is known about the neural mechanisms underlying this component of behavior. Using a unique combination of eye-tracking, neuroimaging, and computer-animated virtual agents, the current investigation demonstrated that gaze-based interactions with a perceived human partner are associated with activity in the ventral striatum, a core component of reward-related neurocircuitry, while interactions with a computer-driven agent activate attention networks. In addition, comparisons of neural activity during interaction with a behaviorally naïve and an explicitly cooperative partner show that that cooperation is not necessary, but that the mere experience of engagement in social interaction is sufficient to recruit the reward system.

One Sentence Summary:

The experience of being engaged in social interaction with another human per se is associated with reward-related activity in the ventral striatum and may provide an explanation as to why we interact with others in the first place.

Main Text:

Introduction

In the hierarchy of human needs, the need to affiliate and interact with others has been located directly after physiological needs and prior to egoistic needs related to self-actualization and esteem (1, 2). This implies that social interaction should represent a reward in itself, and led to the proposal of an intrinsic motivation for social interaction unique to the human species (3–5). Neuroeconomic studies have indeed found reward-related brain activity during social interactions (e.g. 6–9). However, the application of economic games in the study of social interaction typically involves high-level concepts such as trust, fairness, or cooperation (10). As a consequence, the claim that being in interaction with others is per se rewarding has never been put to the test (11).

The current study addressed this notion by studying the brain activity of humans engaged in gaze-based interactions. Gaze constitutes a crucial domain of everyday social encounters (e.g. 12) and has the additional advantage that it can be implemented inside an fMRI scanner due to the minimal involvement of body movements. Ecological validity and experimental control were balanced by visualizing gaze behavior via computer-animated virtual agents (13). Each block of the interaction task comprised five trials in which the agent would either engage in joint or non-joint attention with the participant (Fig. 1A/B). Participants believed that during each block the agent was either controlled by a computer algorithm or another participant. In fact, the other participant was a confederate to permit systematic manipulation of the agent's gaze behavior by the computer algorithm. This was accomplished by varying the proportion of joint attention trials from zero to five out of five possible times (S 3.1). Participants' task was to decide on the nature of their opponent based on the agent's reactions during a given block. While other studies explicated the nature of the opponent a priori (7, 14, 15), the present design assessed the neural mechanisms underlying the subjective experience of social interaction (16).

Unconstrained as well as cooperative interaction contexts were established in two experimental phases in which the interaction partner was either introduced as naïve to participants' task, or as a collaborator motivated to help them to distinguish human interactions. Based on the claim that social interaction is per se rewarding, we hypothesized that an explicitly cooperative context is not required to recruit reward-related neurocircuitry. Furthermore, we

predicted that the striatum would encode reward components related to a motivation to interact, whereas the orbitofrontal cortex was expected to encode the rewarding experience. Finally, the interaction contexts should differentially affect the neural processing of the agent's behavior.

Results

The contingency of the agent's gaze reactions had a differential effect in the two contexts (see S6.1 for results). In the naïve context (Fig. 1C), the proportion of blocks rated as 'human' correlated linearly with the number of joint attention trials per block, which argues for a particular importance of congruent reactions during unconstrained interactions. In the cooperative context (Fig. 1D), 'human' ratings were more closely related to the general contingency of reactions: the engagement in joint as well as non-joint attention was interpreted as indicative of a human opponent when this behavior was consistent over an entire block. This replicates results of a behavioral between-subject version of the present task (16). A series of regression analyses assessed how trial type (joint vs. non-joint) at a given trial position biased participants' rating of an interaction as 'human'. This yielded a measure of the integration of information during decision-making (Table S1, see S6.2 for details). In the naïve context, already the second trial had a significant influence on the final rating, with 'human' becoming significantly more likely than 'computer' if the agent engaged in joint attention on that trial. Such an early component was absent in the cooperative context, where the influence of trial type increased roughly linearly until end of a block.

Initial fMRI analyses (see Table S2 for an overview of analyses) were driven by participants' ratings (Fig. 1C/D). Blocks rated as 'human' (hum_{all_block}>com_{all_block}) were accompanied by enhanced activation of the ventral striatum (VS) and the medial orbitofrontal cortex (mOFC; Fig. 2A, Table S3a). In contrast, during blocks rated as 'computer' (com_{all_block}>hum_{all_block}) there was increased activation of a fronto-parietal attention network (FPAN, see 17) including the inferior parietal cortex, precuneus, and the lateral prefrontal cortex (Fig 2B, Table S3b). When considering only the naïve context ($hum_{naïve\_block}$>$com_{naïve\_block}$), 'human'-rated blocks engaged the mesolimbic reward system (18), while there were no significant results in 'computer'-rated blocks ($com_{naïve\_block}$>$hum_{naïve\_block}$). Conversely, in the cooperative context, there was increased activation of the FPAN during 'computer'-rated blocks

(com_{coop_block}>hum_{coop_block}, Table S3d), while no regions were active during 'human'-rated blocks (hum_{coop_block}>com_{coop_block}).

Further fMRI analyses were driven by the finding of early versus linear integration of gaze reactions in the naïve and cooperative context. Accordingly, VS activity during the first two trials of a block ($hum_{naïve\_first}$>$com_{naïve\_first}$) was predictive of participants' ratings exclusively in the naïve context (Fig. 3B, Table S3e). On the contrary, only in the cooperative context, linear parametric analyses including trial progression as a parametric regressor (hum_{coop_param}>com_{coop_param}) revealed an increase of VS activity over the full length of blocks rated 'human' (Fig. 4, Table S3f).

The processing of single gaze reactions was assessed using event-related analyses, while contingencies established by the agent's gaze reactions over a block (see S3.1) were analyzed using the total number of joint attention trials within a block as a parametric regressor. In the naïve context, joint attention trials ($JA_{naïve}$>$NJA_{naïve}$) recruited regions associated with the 'social brain network' (19) involving the bilateral anterior cingulate cortex (ACC), and the left amygdala, medial prefrontal cortex (mPFC), temporal pole, and superior temporal sulcus (Fig. S2, Table S5a). However, parametric analyses of increasing positive contingency ($naïve_{increase\_JA}$) revealed increases of activity in the bilateral paracentral lobule, but not in social brain regions (Fig. S3A, Table S4a). In the cooperative context, there was no enhanced activity for joint attention on a single-trial level (JA_{coop}>NJA_{coop}), but a parametric increase with increasing positive contingency ($coop_{increase\_JA}$) in the dorsal striatum, the thalamus, the ACC and the mPFC (Fig. S4A, Table S4c), suggesting that the social brain network is recruited by behavioral consistency rather than single events. Increased negative contingency ($coop_{increase\_NJA}$) recruited the FPAN (Fig. S4B, Table S4d).

Discussion

These findings provide first-time evidence that the mere experience of social interaction with another human is sufficient to recruit the mesolimbic reward system (20). A recent study showed a correlation between social reward dependence – i.e. a measure of an individual's propensity to engage in social interaction – and gray matter density in the VS and the mOFC. The structural predisposition for social interaction thus overlaps with brain regions involved in the processing

of primary rewards (4), thereby further lending support to the hypothesis that social interaction constitutes a primary reward.

Notably, the processing of rewards has been divided into 'wanting' and 'liking' components (18). In the naïve context, striatal activity during the first two trials reliably predicted that an interaction will be rated as 'human'. Concordantly, post-hoc ratings revealed that participants largely relied on their intuition in this context (Fig. S1B), and preferred thinking about the behavior of a human conspecific rather than a computer when making the decision (Fig. S1D). As the agent's behavior during the first two trials is actually still inconclusive regarding the nature of the opponent, we presume that this early preference represents the 'wanting' component associated with the human need to interact (1). This is consistent with the previous observation that the VS conveys automatic incentive signals to the mOFC during initial stages of impression formation (21) and thereby contributes to a larger picture attributing the VS to an automatic valuation system which encodes preferences irrespective of stimulus modality and task demands (22).

If a need for interaction represents the 'wanting' component, its fulfillment by the actual experience of being in social interaction should correspond to reward 'liking'. Indeed, participants' post-experience ratings indicated that they experienced social interactions as more pleasant than non-social interactions (Fig. S1C). The subjective hedonic experience of rewards has been attributed to the mOFC rather than the VS (23). The mOFC was active during blocks rated as 'human' irrespective of interaction context. This finding might reflect the 'liking' of being engaged in interaction, and is further supported by the observation that the VS plays different functional roles depending on the context.

In the cooperative context, there was no early activation, but a gradual parametric increase of VS activity with increasing trial progression. Considering that the interaction partner allegedly helped participants in their decision, consistent behavior had to be detected by an accumulation of information over time, rather than by trusting an intuition (16). A recent study examined whether VS activation correlates with such an accumulation of evidence in general, or the accumulation of value in particular (24). In a buying task, positive and negative ratings of a product were sequentially disclosed to participants. The VS and the mOFC specifically updated the representation of value when positive ratings were disclosed, but not generally when novel information was revealed or when information was negative. In our study, each trial of

interactions experienced as social provided positive evidence that the agent is controlled by an actual human cooperator. The differential linear increase of striatal activity with these trials therefore reflects the accumulation of value rather than evidence per se (Fig. 4).

The modulation of gaze processing by interaction context provides further insights into the integration of information underlying reward-based decisions. In the naïve context, activity in social brain regions such as the mPFC and the aSTS was confined to single events of joint attention. These areas are known to be involved in an the inference of other individuals' mental states (12, 25, 26), thereby demonstrating that gaze behavior is processed in a trial-by-trial fashion. In cooperative interactions, single events are only indicative of a human opponent when part of temporally contingent behavior. Accordingly, positive contingency was processed by the dorsal striatum, while negative contingency recruited the FPAN. This interplay of reward and attention networks possibly reflects a distribution of cognitive resources required to ensure that joint attention – a cue that is behaviorally relevant and rewarding (12, 26) – is only considered as a social event when the contingency of reactions is high.

The dorsal striatum has also been recruited in iterated trust-games (8), which require participants to monitor a cooperators' actions across multiple rounds and demand similar temporal binding of positively contingent information as the present task (27). Furthermore, this region plays a role when participants experienced contingency between their actions and a reward, thereby linking the sense of agency to reward-based decisions (28). This is consistent with the recent observation that the initiation of joint attention results in an increased experience of agency (29). The reward value of cooperative interactions hence appears to rely on the establishment of contextually meaningful contingencies between one's own actions and another individual's behavior rather than intuitive valuations of reactions.

In sum, our study provides evidence that the basic sociability of human nature rests upon an urge to interact and a predisposition to experience active participation in social interactions as motivating. While humans might have a predisposition to cooperate, these results suggest that we cooperate for a more basal motive – namely to sustain the interaction with another person.

References and Notes:

1. R. F. Baumeister, M. R. Leary, The need to belong: desire for interpersonal attachments as a fundamental human motivation, Psychol. Bull. **117**, 497–529 (1995).

2. A. H. Maslow, A theory of human motivation, Psych. Rev. **50**, 370–396 (1943).

3. S. Krach, F. M. Paulus, M. Bodden, T. Kircher, The rewarding nature of social interactions, Front. Neurosci. **4**, 22 (2010).

4. M. Lebreton et al., The brain structural disposition to social interaction, Eur. J. Neurosci. **29**, 2247–2252 (2009).

5. M. Tomasello, Why We Cooperate (The MIT Press, Cambridge, MA, 2009).

6. J. Rilling et al., A neural basis for social cooperation, Neuron **35**, 395–405 (2002).

7. A. G. Sanfey, J. K. Rilling, J. A. Aronson, L. E. Nystrom, J. D. Cohen, The Neural Basis of Economic Decision-Making in the Ultimatum Game, Science **300**, 1755–1758 (2003).

8. B. King-Casas et al., Getting to know you: reputation and trust in a two-person economic exchange, Science **308**, 78–83 (2005).

9. K. L. Phan, C. S. Sripada, M. Angstadt, K. McCabe, Reputation for reciprocity engages the brain reward center, Proc. Natl. Acad. Sci. U.S.A **107**, 13099–13104 (2010).

10. E. Fehr, C. F. Camerer, Social neuroeconomics: the neural circuitry of social preferences, Trends Cogn. Sci. (Regul. Ed.) **11**, 419–427 (2007).

11. L. Schilbach et al., Toward a second-person neuroscience, Behav. Brain Sci. (in press).

12. L. Schilbach et al., Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry, J. Cogn. Neurosci. **22**, 2702–2715 (2010).

13. C. J. Bohil, B. Alicea, F. A. Biocca, Virtual reality in neuroscience research and therapy, Nat. Rev. Neurosci. **12**, 752–762 (2011).

14. H. L. Gallagher, A. I. Jack, A. Roepstorff, C. D. Frith, Imaging the intentional stance in a competitive game, Neuroimage **16**, 814–821 (2002).

15. K. McCabe, D. Houser, L. Ryan, V. Smith, T. Trouard, A functional imaging study of cooperation in two-person reciprocal exchange, Proc. Natl. Acad. Sci. U.S.A. **98**, 11832–11835 (2001).

16. U. J. Pfeiffer, B. Timmermans, G. Bente, K. Vogeley, L. Schilbach, A Non-Verbal Turing Test: Differentiating Mind from Machine in Gaze-Based Social Interaction, PLoS ONE **6**, e27591 (2011).

17. M. Corbetta, G. L. Shulman, Control of goal-directed and stimulus-driven attention in the brain, Nat. Rev. Neurosci. **3**, 201–215 (2002).

18. K. C. Berridge, T. E. Robinson, J. W. Aldridge, Dissecting components of reward: "liking", "wanting", and learning, Curr. Opin. Pharmacol. **9**, 65–73 (2009).

19. C. D. Frith, The social brain?, Phil. Trans. R. Soc. B **362**, 671–678 (2007).

20. A. Alcaro, R. Huber, J. Panksepp, Behavioral functions of the mesolimbic dopaminergic system: an affective neuroethological perspective, Brain Res Rev **56**, 283–321 (2007).

21. H. Kim, R. Adolphs, J. P. O'Doherty, S. Shimojo, Temporal isolation of neural processes underlying face preference decisions, Proc. Natl. Acad. Sci. U.S.A. **104**, 18253–18258 (2007).

22. M. Lebreton, S. Jorge, V. Michel, B. Thirion, M. Pessiglione, An Automatic Valuation System in the Human Brain: Evidence from Functional Neuroimaging, Neuron **64**, 431–439 (2009).

23. J. Peters, C. Büchel, Neural representations of subjective reward value, Behav. Brain Res. **213**, 135–141 (2010).

24. S. Gluth, J. Rieskamp, C. Büchel, Deciding When to Decide: Time-Variant Sequential Sampling Models Explain the Emergence of Value-Based Decisions in the Human Brain, J. Neurosci. **32**, 10686–10698 (2012).

25. S. Materna, P. W. Dicke, P. Thier, Dissociable roles of the superior temporal sulcus and the intraparietal sulcus in joint attention: a functional magnetic resonance imaging study, J. Cogn. Neurosci. **20**, 108–119 (2008).

26. E. Redcay, M. Kleiner, R. Saxe, Look at this: the neural correlates of initiating and responding to bids for joint attention, Front. Neurosci. **6**, 169 (2012).

27. K. McCabe, D. Houser, L. Ryan, V. Smith, T. Trouard, A functional imaging study of cooperation in two-person reciprocal exchange, Proc. Natl. Acad. Sci. U.S.A. **98**, 11832–11835 (2001).

28. E. M. Tricomi, M. R. Delgado, J. A. Fiez, Modulation of Caudate Activity by Action Contingency, Neuron **41**, 281–292 (2004).

29. U. Pfeiffer et al., Eyes on the Mind: Investigating the Influence of Gaze Dynamics on the Perception of Others in Real-Time Social Interaction, Front. Psychol. **3**, 537 (2012).

Acknowledgments:

The authors would like to thank the members of the Neuroimaging Group at University Hospital Cologne for helpful comments on the manuscript. Furthermore, the authors are grateful to Dorothé Krug and Barbara Elghahwagi for their valuable assistance in data acquisition. This study was supported by a grant of the Köln Fortune Program of the Medical Faculty at the University of Cologne to LS and by a grant "Other Minds" of the German Ministry of Research and Education to KV.

Figures

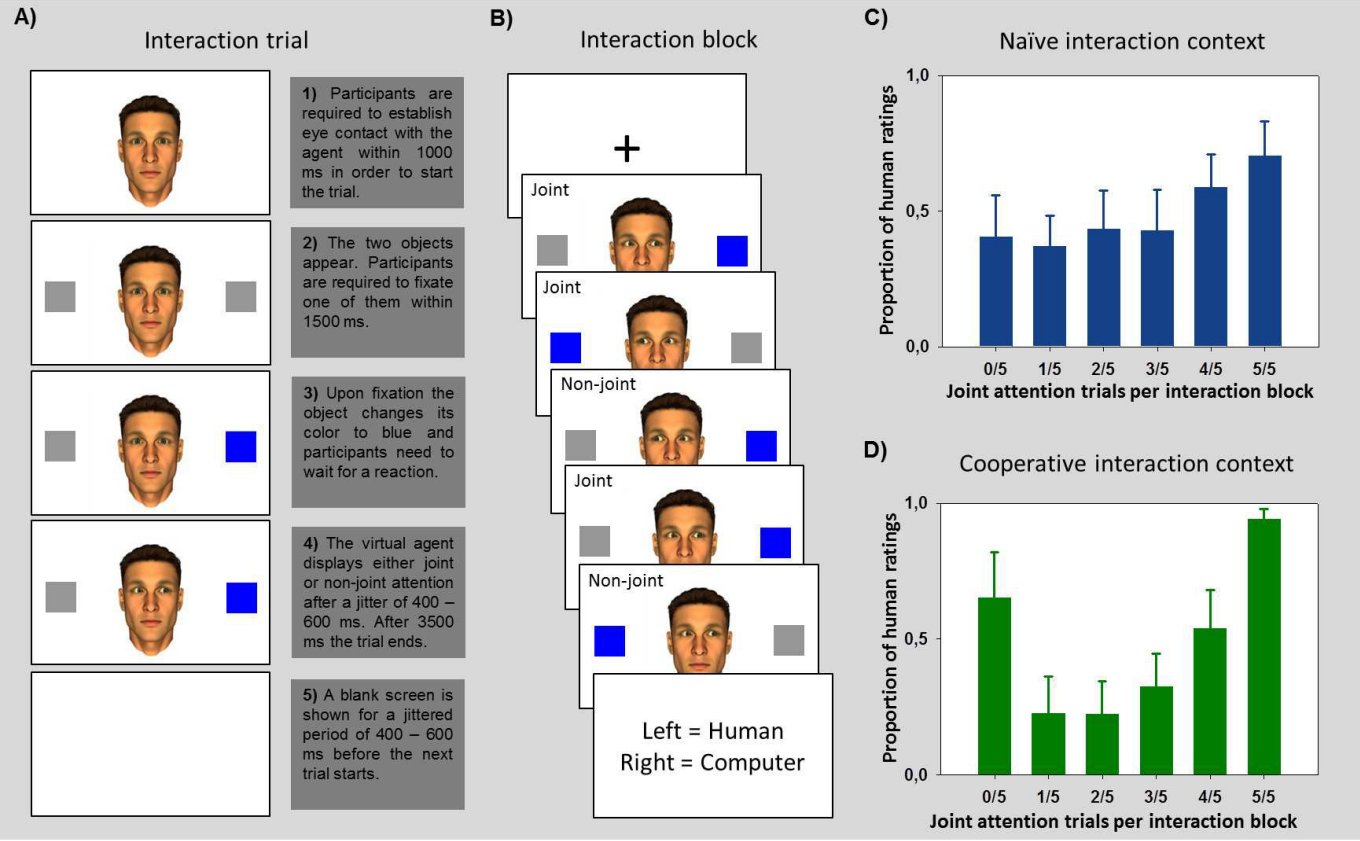

Fig. 1. Structure of interaction task and behavioral results. A) In each of five trials of an interaction block, participants initiate an exchange of gaze shifts. The detailed sequence of events is described in the figure. B) At the end of each block participants indicate whether they experienced this interaction as social ('human') or non-social ('computer'). This block exemplifies a 3/5 condition in which the agent engages three out of five times in joint attention. C) In the naïve context, the proportion of 'human' ratings correlates with increased congruency of gaze reactions. D) In the cooperative context, the mere contingency of gaze reactions modulates the experience of social interaction.

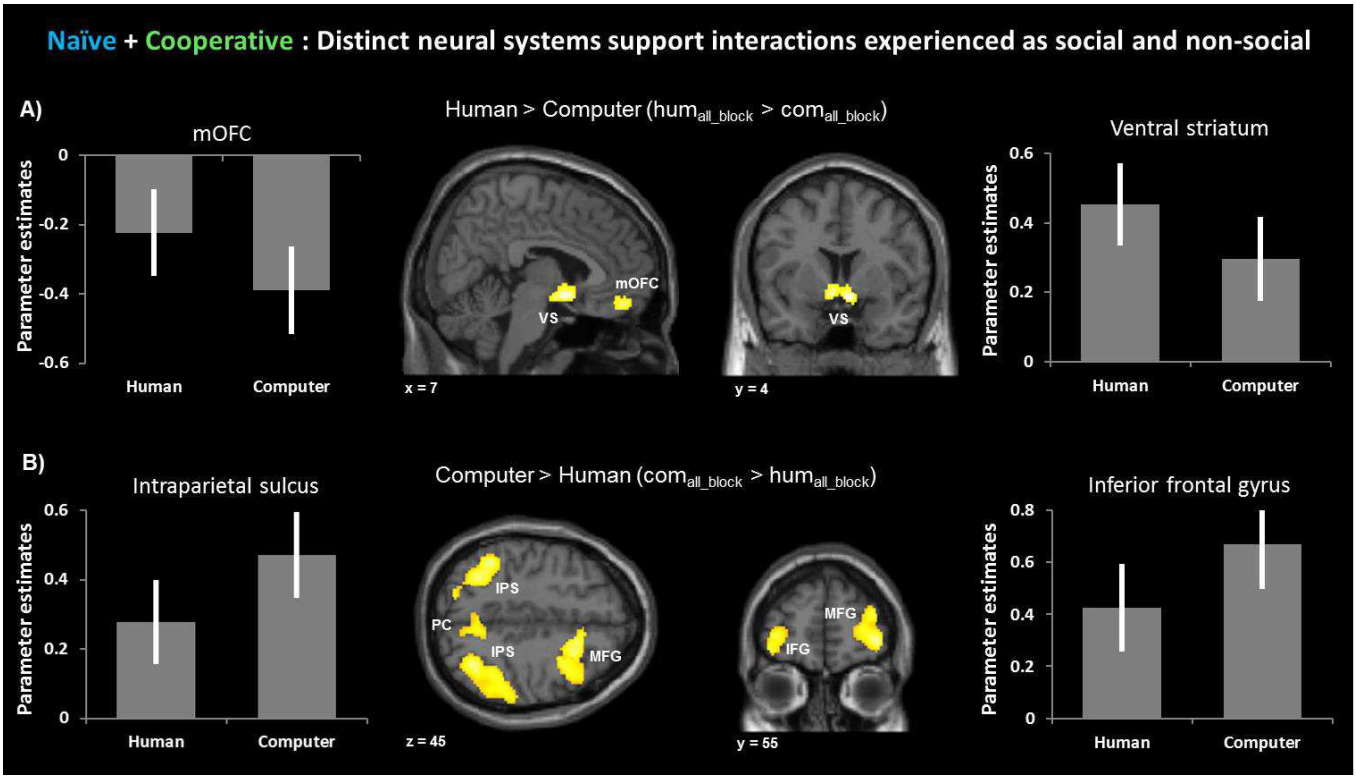

Fig. 2. Neural systems activated during the blocks rated as 'human' and 'computer' independent of the interaction context. A) The experience of interaction with another human participant recruits the ventral striatum (VS) and medial orbitofrontal cortex (mOFC). B) The experience of an interaction as non-social is associated with activity in a fronto-parietal network related to attentional processing, which includes the inferior parietal sulcus (IPS), precuneus (PC), premotor cortex (PMC), and lateral prefrontal cortex including IFG and MFG (statistical threshold: $p < .05$ corrected for multiple comparisons at the cluster level).

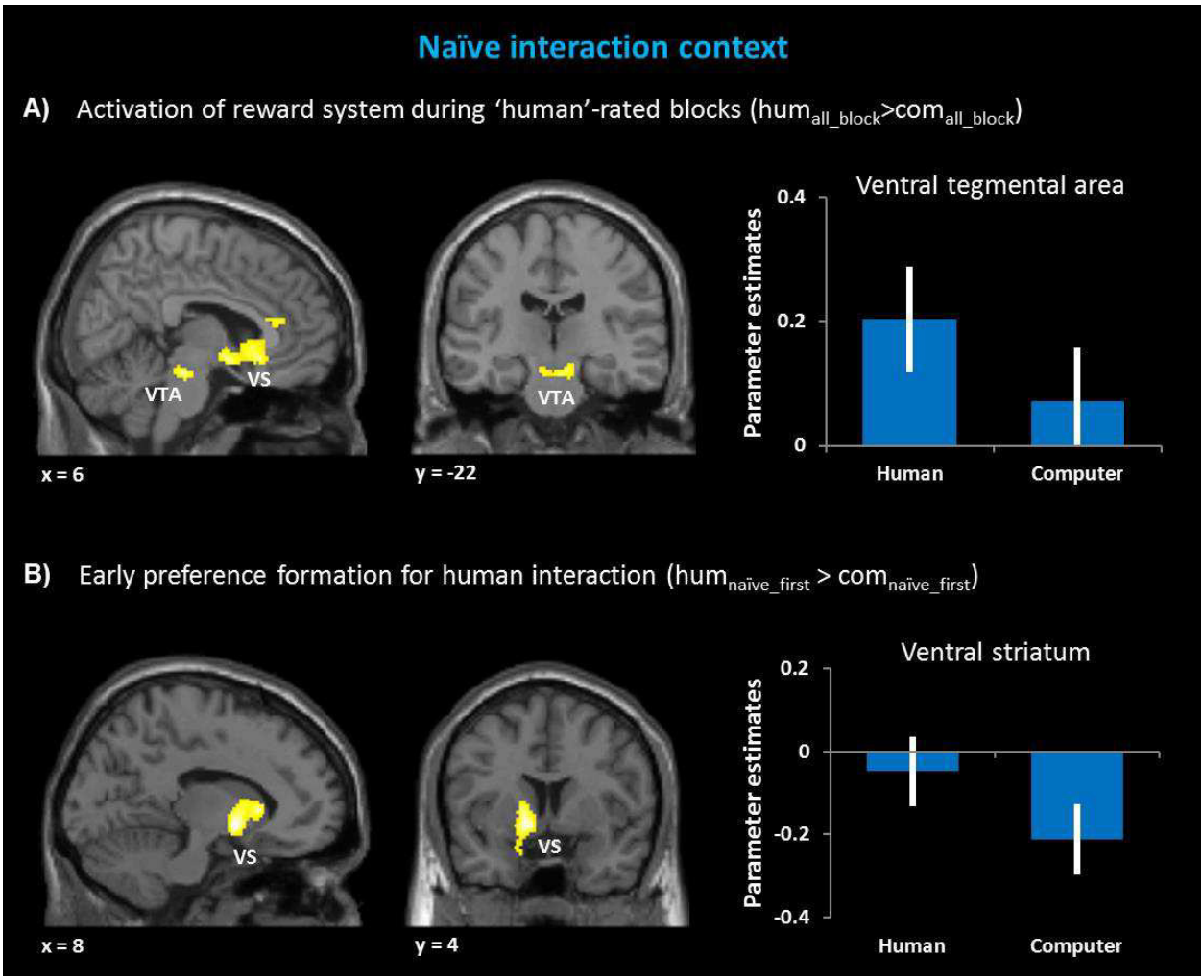

Fig. 3. In the naïve context, the reward system is activated during the experience of engagement in social interaction. A) Blocks rated as 'human' recruit the mesolimbic reward system relative to 'computer'-rated blocks. This includes the ventral tegmental area (VTA), substantia nigra (SN), subthalamic nucleus (STN), and the ventral striatum (VS). B) The activity of the VS during the first two trials of a block is predictive of participants' ratings. This finding parallels behavioral data showing that there is an early critical component in the decision-making process exclusively in the naïve context (threshold: $p < .05$ cluster-level-corrected for multiple comparisons; error bars depict 90% CI).

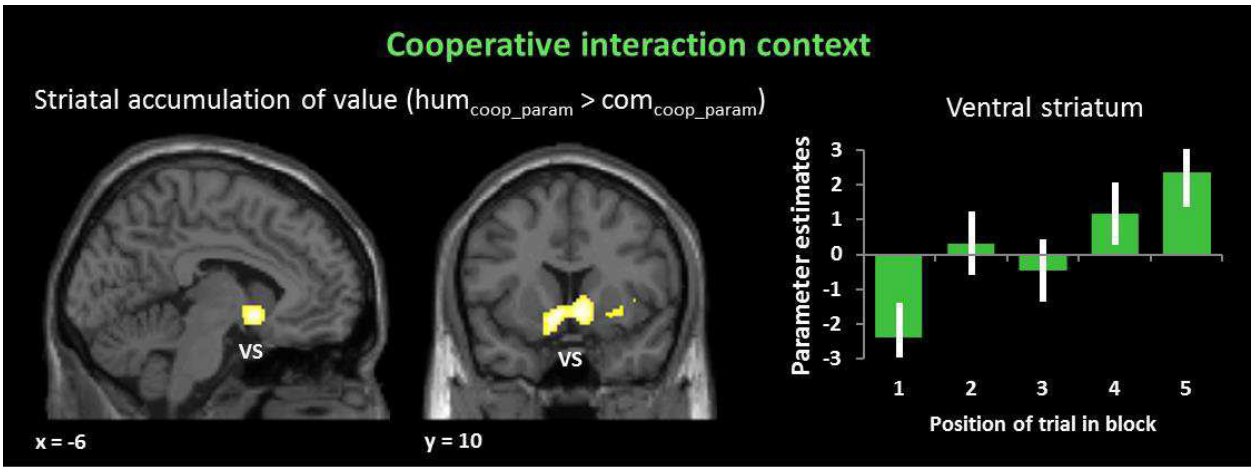

Fig. 4. During interaction with a cooperative interaction partner, activity in the ventral striatum (VS) unfolds over the time course of interaction blocks rated as 'human'. This indicates an accumulation of value towards the decision (threshold: $p < .05$ cluster-level-corrected for multiple comparisons; error bars depict 90% CI).

Supplementary Materials

List of supplementary materials

Materials and Methods

S1 Participants

In total, a group of 32 healthy right-handed volunteers without any records of neurological or psychiatric illnesses participated in the study. All participants had normal or corrected-to-normal vision. They were naïve with respect to the task and were equally compensated for their participation (15 Euro/hour). All participants signed a written consent form in which they confirmed that they participated voluntarily and approved that data are used in an anonymized fashion for analysis and publication. The study followed the WMA Declaration of Helsinki (Ethical Principles for Medical Research Involving Human Subjects) and was approved by the ethics committee of the Medical Faculty of the University of Cologne, Germany. Four participants had to be excluded due to excessive head movements causing uncorrectable signal spiking. Another six participants were dismissed due to technical problems with the eye-tracker which resulted in data due to invalid trials greater than 25 percent. Finally, two participants had to be excluded as they did not believe the cover story. Consequently, 20 participants (9 female/11 male, M = 27.75, SD = 6.44, age range from 21 to 42 years) were included in the analysis.

S2 Materials

S2.1 Stimulus presentation and eye-tracking

Visual stimuli were presented to participants using a custom-built and shielded TFT screen attached at a distance of about 100 cm from the end of the scanner (viewing angle: 14° x 18° horizontal x vertical). They were displayed to participants via a mirror that was mounted on the head coil. Participants' eye-movements were monitored via the same mirror system using the MRI-compatible long-range mount version of the EyeLink 1000 eye-tracking system (SR Research, Missisauga, Canada) installed at the rear end of the scanner bore. Raw gaze data were collected at a sampling rate of 500 Hz on a dedicated EyeLink host computer and made available

to the software package Presentation™ (Version 14.9, Neurobehavioral Systems, Albany, CA) to control stimulus presentation in a gaze-contingent fashion using an interactive eye-tracking algorithm (for details see 30).

S3 Experimental design

The paradigm used here is an fMRI adaptation of the 'non-verbal Turing test' which has recently been validated in a behavioral study as a tool allowing the separation of the experience of social and non-social interaction (16). Participants' task was to engage in gaze-based interaction with an anthropomorphic virtual agent with a neutral facial expression in brief interaction blocks (see supplementary text S7 for a discussion of the use of virtual agents). To ensure comparability of results, the same agent was used as in previous studies (12, 16, 29, 31).

S3.1 The interaction task

Participants had to interact with a virtual agent in a series of interaction blocks. They were instructed that in each given interaction block the gaze reactions of the agent could either be controlled by a real person or a computer algorithm. Based on the gaze behavior of the agent, their task was to indicate whether it had been controlled by another human participant or a computer program at the end of each block. In fact, the interaction partner was a confederate of the experimenter and the gaze reactions of the agent were always controlled by an algorithm. This allowed for systematic variation of gaze reactions. Each interaction block comprised five gaze trials (Fig. 1A/B) in which the agent would either engage in joint or non-joint attention with the participant (see S8 for a discussion of joint attention). Systematic manipulation of the agent's gaze reactions therefore resulted in six experimental conditions (0/5, 1/5, 2/5, 3/5, 4/5, and 5/5 times of joint attention). In addition, a 'high-level baseline' in which the agent closed its eyes during each trial of a block was modeled separately and served as a control condition. A similar control condition has been used in other studies on joint attention (26, 32).

In sum, participants' ratings provided a way to experimentally distinguish the subjective experience of being engaged in social interaction with another person from a non-social form of interaction – i.e. an interaction with a computer-animated virtual agent. Participants' decision

thereby was a dependent, instead of an independent variable as in other studies employing a human-computer distinction (7, 14, 15). Furthermore, the systematic manipulation of the agent's gaze behavior on a trial-by-trial basis resulted in an equally systematic manipulation of the agent's behavioral contingency over time. Contingency hereby refers to a consistent causality between one's own actions and another person's reactions during a given interaction block. In the present design, behavioral contingency increases the more often an agent displays the same reaction to a participant's gaze shift during one interaction block, irrespective of the nature of this reaction. This means that an agent engaging in joint attention in each trial of a block behaves as contingent as an agent engaging in non-joint attention in each trial. The increase of the former is referred to as positive congruency (i.e. congruency), whereas the increase of the latter is referred to as negative contingency.

S3.2 Structure of gaze-based interaction blocks

Each trial followed a particular sequence (Fig. 1A depicts an exemplary trial): First, participants were required to look at a virtual agent appearing on the screen within 1000 milliseconds. Once the algorithm had detected a fixation of the agent, two grey squares appeared on the left and the right side of the screen. Subsequently, participants had to choose one of the squares by fixating it. In case the agent was not fixated within 1500 ms following the start of the trial, or in the absence of any successful object fixation within 1500 ms, the trial was aborted and the next trial of the interaction block began. Blocks containing aborted trials were registered as invalid. Upon successful fixation, the selected object was marked in blue in order to provide participants with feedback about successful registration of their gaze. Participants were informed that their initial gaze shift to the square (but not the color change) was transmitted to the screen of the other participant in real-time and that they would see the other participant's response to their gaze shift as visualized by the gaze of the agent on their screen. With a latency jittered between 400 and 600 ms which has previously been demonstrated as 'natural' for human gaze reactions (29), the agent would then either follow the participant's gaze to the selected object or avert its gaze to the other object. This resulted in situations of joint and non-joint attention, respectively (S8). The resulting screen configuration was displayed to participants until the trial ended after 3500 ms. Next, a blank screen was shown for 400 – 600 ms before the next trial started. This was signaled by the re-appearance of the agent's face. The duration of the whole trial amounted to 4000 ms.

After each block of five trials, participants were asked to indicate whether they believed they had been interacting with the other human participant or a computer algorithm (i.e. the dependent variable). They did so by pressing one of two buttons with their index fingers within a response window of 1500 ms. The side of the buttons for the two options was balanced between participants. The response window was followed by a fixation cross that was shown for a jittered period of 5000 to 7000 ms until the next block started. Blocks in which responses were missing were marked as invalid.

S3.3 Naïve and cooperative interaction contexts

The experiment consisted of two phases to distinguish unconstrained from cooperative interactions. In the first phase, the confederate was introduced as naïve to the participant's task. Participants were told that their interaction partner had been instructed to react to each of their gaze shifts by looking at one of the two objects without any additional information. They were explicitly informed that their interaction partner thus could not willingly help them in their decision. This condition was included to provide a spontaneous and unconstrained interaction context in which participants had no a priori assumptions about their interaction partner's behavior. In the second phase, the other participant was introduced as explicitly cooperative. To this end, participant were instructed that the interaction partner had been debriefed about the participant's task and the involvement of a computer algorithm randomly taking control over the agent's gaze reactions in certain interaction blocks. Participants were further told that the interaction partner's task during the second phase of the experiment was to behave explicitly cooperative, that is, to react to their gaze shifts in such a way that would make it as easy as possible for them to distinguish between human and computer interaction. It was not specified in which exact way the interaction partner was supposed to achieve this. Each phase of the experiment consisted of two runs (see S3.4). During each run, the six experimental conditions and the control condition were repeated three times in a randomized fashion. Each gaze condition was thus repeated six times during the naïve as well as during the cooperative phase.

The order of the two experimental phases could not be randomized, as the naïve condition required participants to assume that their interaction partner did not have any knowledge about the nature of their task and reacted entirely and spontaneously based on their personal intentions. However, in order to control for novelty and habituation effects, participants

engaged in a practice session of 5 minutes before the start of the first run. In addition, participants received a detailed written explanation of the task and were able to ask questions after the practice session in case any aspect of the procedure was not clear to them. The experimenter closely monitored their eye-movements and behavior during the practice session to provide additional instruction, if necessary. As noted above, the design was a within-subject adaptation of a previous behavioral study. Behavioral results are an exact repetition of the results of the between-subject version in which different groups of participants interacted with naïve and cooperative partners (16). Furthermore, participants' responses in the debriefing questionnaires did not indicate any differences between the within- and between-subject version (compare Figures S1 of the present study and 16). For these reasons, we are convinced that potential sequence effects are negligible.

S3.4 Procedure

Upon participants' arrival at the MRI facility, they were asked to take a seat in a room dedicated to the preparation of participants. It was disclosed to them that they would interact with another human participant outside the scanner in real-time by moving their eyes. The application of computer-animated virtual characters for the visualization of gaze was explained as being due to technical constraints. After the initial cover story, they received written instructions covering the exact procedure of the task. Upon reading these instructions, they were given the opportunity to ask questions for clarification of the instructions. Subsequently, the experimenter walked them past another room, where a confederate was sitting who was allegedly starting to read her instructions in that very moment. This person was quickly introduced as the interaction partner who was purportedly being instructed while the actual participant was prepared for MRI in the scanner room. It was made sure that participants only met for few seconds without being able to exchange more than a brief greeting. The participant was then led to the scanner room and prepared for the experiment. After the eye-tracker was calibrated, participants were informed that their interaction partner was now instructed about her task and ready to begin.

The scanning session then started with a practice session consisting of a series of eight interaction blocks to acquaint subjects with the interaction paradigm. After the practice session, participants had the opportunity to ask questions for clarification if necessary. Eventually, the experiment started with the first phase, in which the interaction partner was introduced as naïve

to the task of the participants. After the first two runs of the experiment, there was a break of about 3 minutes before the second phase began (see S3.3) which again consisted of two experimental runs. Each run lasted about 10 minutes. Total scanning time including the practice session and short breaks between the runs therefore amounted 50 minutes. After the experiment, participants were led back to the preparation room where they were asked to fill out a post-experiment questionnaire (Fig. S1). After data collection was completed, all participants received an email debriefing them about the deception with respect to the confederate. This email contained detailed information as to why this deception was necessary and explicated the purpose as well as preliminary results.

S4 Functional magnetic resonance imaging

S4.1 Data acquisition

Scanning was performed on a Siemens Trio 3-T whole-body scanner (Siemens Medical Solutions, Erlangen, Germany). For the fMRI scans a T2*-weighted gradient echo planar imaging (EPI) sequence with the following parameters was used: TR = 2200 ms, TE = 30 ms, 36 axial slices, slice thickness = 3.0 mm, in-plane resolution = 3.0 x 3.0 mm, field of view = 200.0 x 200.0 mm. In each session, 280 images were acquired. The first five images of a session were discarded before image processing in order to eliminate potential saturation effects.

S4.2 Preprocessing

Images were analyzed using SPM8 (Wellcome Department of Imaging Neuroscience, London, UK) implemented in MATLAB 7.1 (Mathworks Inc., Natick, USA). DICOM images were converted to NIFTI format before further image processing. Preprocessing involved the following steps: Motion correction was completed by an affine registration procedure (33). Realignment was performed in two steps. Images were initially realigned to the first image of the time series and subsequently to the mean of these images. For normalization, the mean EPI was computed for each participant and spatially normalized to the MNI single subject template (34) using the unified segmentation function of SPM8 with a 2x2x2 mm isotropic resolution. The

ensuing deformation was then applied to the individual EPI volumes. Finally, spatial smoothing of the normalized images was performed using an 8-mm FWHM Gaussian kernel.

S5.2 Data analysis strategies

The design of the study justified three lines of data analysis that are explained in the following. For a comprehensive overview of all analyses and contrasts see Table S2.

(1) Analyses based on the dependent variable: Initially, data analysis was guided by participants' ratings to examine differences in neural activity elicited by the experience of social and non-social interaction, that is, by 'human' (hum) and 'computer' (com) ratings, respectively [contrasts of interest: ($hum_{all\_block} > com_{all\_block}$), ($com_{all\_block} > hum_{all\_block}$)]. Stimulus events were defined by the onsets of interaction blocks and their duration until the start of the button press response window (20000 ms). If present, invalid blocks were modeled with a distinct regressor of no interest. Subsequently, the effect of the instruction was investigated by conducting separate analyses for the naïve and cooperative interaction contexts. Blocks rated as 'human' and 'computer' were modeled separately for the two contexts [contrasts of interest: ($hum_{naïve\_block} > com_{naïve\_block}$), ($com_{naïve\_block} > hum_{naïve\_block}$), ($hum_{coop\_block} > com_{coop\_block}$), ($com_{coop\_block} > hum_{coop\_block}$)]. Two further analyses were informed by the output of the logistic regressions performed on the behavioral data which are described in section S6.2 of the supplementary text. Results of these sequential regressions demonstrated that during naïve interaction there was an early effect of trial type on the final decision, whereas during cooperative interaction the increase of the importance of trial type was roughly linear over the course of interaction blocks experience as social. As a consequence, one neural analysis aimed at investigating early components in the experience of social interaction by confining a stimulus event to the first two trials (8000 ms) of an interaction block [contrasts of interest: ($hum_{naïve\_early} > com_{naïve\_early}$), ($com_{naïve\_early} > hum_{naïve\_early}$), ($hum_{coop\_early} > com_{coop\_early}$), ($com_{coop\_first} > hum_{coop\_first}$)]. In the second analysis, trial progression was introduced as a parametric regressor to model increases of neural activity with increasing trial position towards the decision [contrasts of interest: ($hum_{naïve\_param} > com_{naïve\_param}$), ($com_{naïve\_param} > hum_{naïve\_param}$), ($hum_{coop\_param} > com_{coop\_param}$), ($com_{coop\_param} > hum_{coop\_param}$)].

(2) Analyses based on the independent variable: As an independent variable, the contingency of gaze behavior was manipulated from zero to five out of five possible occurrences

of joint attention, thereby resulting in six experimental conditions. Driven by this manipulation, a combined categorical-parametric analysis was conducted to model increases and decreases of neural activity with increasing and decreasing numbers of joint attention trials (i.e. positive and negative contingency) per block. At the subject-level, the six experimental categories (0/5, 1/5, 2/5, 3/5, 4/5, and 5/5 possible joint attention trials) were modeled as separate contrasts. In addition, there was a control condition in which the agent closed its eyes during each trial, which was modeled with a separate regressor. The effect of the increasing number of joint attention trials was then modeled over these categories as a linear parametric modulation of the hemodynamic response [contrasts of interest: ($naïve_{increase\_JA}$), ($coop_{increase\_JA}$)]. Accordingly, the effect of increasing numbers of non-joint attention trials (NJA) was modeled by an inverted linear parametric modulation [contrasts of interest: ($naïve_{increase\_NJA}$), ($coop_{increase\_NJA}$)].

(3) Event-related analysis of single gaze trials: Finally, using event-related analyses we compared how JA and NJA trials were processed in the naïve and cooperative condition in order to further investigate the integration of single gaze trials. To this end, joint and non-joint attention trials were modeled as separate regressor. This was done irrespective of the experimental category (0/5, 1/5, 2/5, 3/5, 4/5, 5/5) a trial was presented in [contrasts of interest: ($JA_{naïve}$>$NJA_{naïve}$), ($NJA_{naïve}$>$JA_{naïve}$), (JA_{coop}>NJA_{coop}), (NJA_{coop}>JA_{coop})]. Stimulus events were defined from the onset of a trial until the appearance of the blank screen, thereby amounting to a trial length of 3500 ms (Fig. 1A). Finally, the temporal derivatives of the hemodynamic response function were included in the model (35).

S5.3 Statistics

Data were analyzed using a General Linear Model as implemented in SPM8. Low-frequency signal drifts were removed from the fMRI time series using a high-pass filter with a cutoff of 128 seconds (36). In each of the block-level analyses, experimental conditions were modeled by a boxcar reference vector which was convolved with the canonical hemodynamic response function. At the group-level, individual contrasts obtained from subject-level analyses were fed to a flexible factorial ANOVA with factors subject and condition using a random-effects model (37). Differences between conditions and deviations from zero were investigated using linear contrasts of the group-level parameter estimates. All effects were thresholded at $p < .05$ at the cluster-level, family-wise-error-corrected for multiple comparisons ($p_{FWE\text{-}corr} < .05$), with an

underlying voxel-level threshold of $p < .001$, uncorrected. Anatomical localization of activations was achieved by using version 1.8 of the SPM anatomy toolbox (38) and the brain atlas of Duvernoy (39). Activation maps were superimposed on an SPM canonical T1-weighted image.

Supplementary text

S6 Behavioral results

S6.1 Effect of gaze condition and instruction on the experience of social interaction

The effect of interaction context (naïve vs. cooperative) and gaze contingency (total number of joint attention trials per interaction block) on participants' subjective experience of the interaction (i.e. their ratings) was analyzed using repeated-measures ANOVAs. Planned polynomial contrasts were applied for trend analysis to describe behavioral patterns statistically. Prior to data analysis, an arcsine transformation was performed as proportional data violate the assumption of normality (40). The Greenhouse-Geisser correction was applied in all analyses.

Fig. 1 depicts the proportion of blocks rated as 'human' (i.e. the dependent variable) depending on the contingency of gaze reactions (i.e. the independent variable) during the naïve (Fig. 1c) and the cooperative condition (Fig. 1d). Results demonstrated a main effect of gaze on participants' ratings, $F(2.45, 46.55) = 13.19$, $p < .001$, $\omega^2 = .23$, and a significant interaction between instruction and contingency, $F(3.13, 59.35) = 11.19$, $p < .001$, $\omega^2 = .08$. This interaction was scrutinized by analyzing the results of the naïve and the cooperative condition separately. In the naïve condition, there was a significant main effect of the factor contingency on participants' ratings, $F(2.76, 52.38) = 3.55$, $p = .023$, $\omega^2 = .03$. Planned polynomial contrasts revealed that this effect was characterized by a significant linear trend, $F(1, 19) = 7.84$, $p = .011$, $\omega^2 = .29$, thereby indicating that the proportion of blocks rated as 'human' increased with increasing numbers of joint attention trials per interaction block (Figure 1B). In the cooperative condition, contingency also had a significant effect on participants' ratings, $F(2.79, 52.95) = 21.79$, $p < .001$, $\omega^2 = .27$. This was characterized by a significant linear trend, $F(1, 19) = 20.21$, $p < .001$, $\omega^2 = .19$, and a quadratic trend, $F(1, 19) = 36.63$, $p < .001$, $\omega^2 = .39$. These trends confirm that when interacting with a cooperative partner, the ascription of humanness is driven not only by congruency but also by contingency of gaze reactions (Fig. 1C).

S6.2 Influence of trial type

Furthermore, to get more insight into how participants' decision-making process unfolded over an interaction block, we assessed how ratings were influenced by consecutive trials. We first performed a regression analysis that included ten predictors: five Trial predictors for the agent's reactions on each of the five trials, and five Trial x Instruction predictors, to look at whether there were any differences in trial influence between instruction conditions (naïve vs. cooperative). In order to focus exclusively on the important factors within the set of ten, we used a Forward method, in which predictors are added consecutively, starting with the strongest, and predictors with decreasing strength are added to the model until adding a new predictor does fail to explain significantly more variance. The results are listed in the top half of Table S1. The final three trials show up as main effects across both conditions, with the fourth trial weighing in the heaviest: if on this trial the agent follows the participant's gaze, the chances of rating 'human' are over 1.59 times higher than when the agent looks the other way. Most importantly however, the second trial loads only in interaction with condition. We therefore performed separate Forward method regression analyses for each interaction context (naïve vs. cooperative), in which the five predictors were the agent's reactions on the five trials (bottom of Table S1). Apart from confirming the persistence of the main effects for the final three trials, the analyses show that only in the naive condition there is an early component in the decision making process that is completely absent in the cooperative condition: the second trial has in fact the second biggest influence on the eventual humanness rating, with 'human' becoming 1.38 times more likely if this second trial consists of JA. In the cooperative condition, the second trial doesn't load at all. In sum, there is an early influence of trial type on participants' ratings in the naïve context, which is absent in the cooperative context, where the integration of information related to decision-making is roughly linear.

S7 Interaction with virtual agents

Methods borrowing from virtual reality technology have found their way into numerous areas of neuroscientific research (13). Specifically in social psychology, the use of virtual agents has been promoted as they allow the isolation of a cue of interest, such as gaze reactions, while

neutralizing confounding variables (41). This becomes particularly evident in neuroimaging studies, which require a substantial reduction of the 'band width' of social interactions in order to maintain experimental control. Recent research on so-called anthropomorphic virtual agents (i.e. as used in this study) - which have realistic human features while still being easily recognized as artificial - has suggested that they provide an excellent tool to study social cognition face-to-face and in real-time (42). More precisely, the interaction with anthropomorphic agents results in comparable reactions and social behaviors as the interaction with real humans. For example, socially induced inhibition as well as facilitation of task performance could not only be observed in the presence of another person, but also in the presence of human-controlled avatars (43). Likewise, participants' regulation of interpersonal distance and approach behavior is comparable in immersive virtual environments and real social encounters (44). It was also demonstrated that participants displayed empathic concern for virtual agents in distressful situations, thereby demonstrating that encounters with virtual agents readily elicit prosocial behavior (45). Finally, a recent study showed that the simulation of gaze behavior by virtual agents in face-to-face interactions results in similar experiences of social presence and intimateness as real gaze behavior (46). In sum, the use of virtual agents enables a balance between ecological validity and experimental control and is thus ideally suited for the purpose of studying social interactions.

S8 Joint attention and social interaction

In the present task, social interactions have been operationalized by virtual agents who engage to different degrees in joint (JA) or non-joint attention (e.g. NJA) with participants. It shall hence briefly be explained here why JA is used to implement social interactions. JA represents a triadic interaction involving a 'referential triangle' of two individuals and some third entity in the environment (47). Considering that people look where they attend and where they intend to act, JA is considered fundamental to an understanding of other minds. An important distinction is made between responding to other people's offerings of JA and initiating JA (48). While responding to a bid for joint attention by following someone's gaze can also be observed in non-human primates, the initiation of JA represents a uniquely human capacity. Accordingly, it has been demonstrated that starting at the age of 9 months, children show an intrinsic motivation to

actively share attention with parents and caretakers by directing their gaze to an object, and to reengage them in social interactions which have previously been interrupted (see 49 and references therein). JA thereby combines an intrinsic motivation to share with the establishment of perceptual common ground in a reciprocal fashion (50). This has been argued to create "a shared space of common psychological ground that enables everything from collaborative activities with shared goals to human-style cooperative communication" (51, p. 121). In sum, JA can be considered as an origin of any meaningful non-verbal social interaction. This suggests that constructing interactions by manipulating the engagement in JA allows studying engagement and participation in real-time social interactions in a socially salient and ecologically valid fashion.

Supplementary Tables

Table S1. Regression coefficients for the logistic regression models (Forward method) for each condition separately, and for both conditions with interaction effects.

	Coeff	SE	Wald	Odds ratio	95 % CI
NAIVE + COOP: Main effect of trials and interaction of trials*condition					
2nd trial (NJA vs JA) * Condition (NAIVE vs COOP)	0.284	0.087	10.70**	1.33	[1.12 - 1.58]
3rd trial NJA vs JA	0.293	0.079	13.63***	1.34	[1.15 - 1.57]
4th trial NJA vs JA	0.464	0.08	34.08***	1.59	[1.36 - 1.86]
5th trial NJA vs JA	0.309	0.08	15.02***	1.36	[1.17 - 1.59]
NAIVE: Main effect of trials					
2nd trial NJA vs JA	0.32	0.113	8.07**	1.38	[1.10 - 1.72]
3rd trial NJA vs JA	0.255	0.114	4.98*	1.29	[1.03 - 1.62]
4th trial NJA vs JA	0.432	0.114	14.32***	1.54	[1.23 - 1.93]
5th trial NJA vs JA	0.264	0.115	5.31*	1.3	[1.04 - 1.63]
COOP: Main effect of trials					
3rd trial NJA vs JA	0.328	0.111	8.76**	1.39	[1.12 - 1.72]
4th trial NJA vs JA	0.489	0.112	19.09***	1.63	[1.31 - 2.03]
5th trial NJA vs JA	0.346	0.112	9.52**	1.41	[1.14 - 1.76]

Notes: *** p<.001; ** p<.005; * p<.05.

Table S2. Overview of the different fMRI analyses with references to the respective results tables.

	All	Naive	Coop
Analysis by dependent variable			
Whole block			
Human > Computer	Table S3a	Table S3c	n.s.
Computer > Human	Table S3b	n.s.	Table S3d
First two trials			
Human > Computer	-	Table S3e	n.s.
Computer > Human	-	n.s.	n.s.
Parametric increase over trials			
Human > Computer	-	n.s.	Table S3f
Computer > Human	-	n.s.	Table S3g
Analysis by independent variable			
Parametric increase (pos. contingency)	-	Table S4a	Table S4c
Parametric decrease (neg. contingency)	-	Table S4b	Table S4d
Event-related analysis			
JA > NJA	-	Table S5a	n.s.
NJA > JA	-	n.s.	Table S5b

Table S3. Analyses based on participants' subjective ratings reveal brain regions differentially activated during 'human'- and 'computer'-rated blocks.

Region	Cluster		Side	MNI Coordinates			T
	Size	$p_{FWE\text{-}corr}$		*x*	*y*	*z*	
Whole Block							
a) $hum_{all\_block} > com_{all\_block}$							
Ventral striatum	468	.000	R	8	4	-10	4.96
Ventral striatum			L	-6	10	-8	4.57
mOFC	326	.002	R	4	48	-16	4.61
mOFC			L	-8	42	-12	3.98
b) $com_{all\_block} > hum_{all\_block}$							
Supramarginal gyrus	2919	.000	R	52	-42	36	4.88
Intraparietal sulcus			R	34	-60	44	4.88
Precuneus			R	10	-60	40	4.67
Intraparietal sulcus	1513	.000	L	-30	-48	32	5.63
Inferior frontal gyrus	1455	.000	R	40	56	-2	5.04
Lateral orbital sulcus			R	46	48	-12	4.93
Middle frontal gyrus			R	52	38	22	4.39
Middle frontal gyrus	1207	.000	R	32	6	60	4.89
Superior frontal sulcus			R	22	14	42	4.82
Inferior frontal gyrus	332	.013	L	-36	56	2	4.39
Lateral orbital gyrus			L	-40	54	-8	4.21
c) $hum_{naïve\_block} > com_{naïve\_block}$							
Ventral striatum	836	.000	R	10	4	-10	5.17
Medial orbitofrontal cortex			R	6	22	-10	4.40
Putamen			R	24	20	0	4.14
Anterior cingulate cortex	299	.020	L	-8	32	6	4.54
Anterior cingulate cortex			R	10	28	16	3.75
Ventral striatum	248	.039	L	-10	2	-2	4.19
Putamen			L	-8	16	2	4.14
Medial orbitofrontal cortex			L	-8	28	-12	3.98
Substantia nigra/Subthalamic nucleus	243	.043	L	-6	-6	-16	4.53
Ventral tegmental area			R	6	-24	-18	4.46

Ventral tegmental area			L	-6	-20	-18	4.07
d) com_{coop_block} > hum_{coop_block}							
Intraparietal sulcus	10451	.000	R	36	-44	34	6.23
Precuneus			R	8	-58	42	6.11
Supramarginal gyrus			R	50	-44	36	6.03
Intraparietal sulcus			L	-32	-54	40	5.87
Inferior frontal sulcus	6743	.000	R	28	52	4	6.56
Inferior frontal gyrus			R	40	56	-2	6.29
Middle frontal gyrus			R	32	8	60	6.21
Middle frontal gyrus	936	.000	L	-50	24	34	6.47
Inferior frontal sulcus			L	-28	16	32	5.03
Lateral orbital gyrus	868	.000	L	-35	56	-8	5.17
Inferior frontal gyrus			L	-34	48	2	4.98
Middle frontal gyrus	452	.003	L	-30	0	60	4.76
First two trials of a block							
e) $hum_{naïve\_early}$ > $com_{naïve\_early}$							
Nucleus accumbens (Ventral striatum)	704	.012	L	-12	6	-6	4.35
Caudate nucleus (head)			L	-10	18	2	4.08
Putamen			L	-18	18	-8	3.92
Medial orbital gyrus			L	-22	14	16	3.93
Parametric increase over whole block							
f) hum_{coop_param} > com_{coop_param}							
Nucleus accumbens (Ventral striatum)	719	.000	R	6	6	-4	4.49
Nucleus accumbens (Ventral striatum)			L	-6	10	-8	4.45
Putamen			R	20	14	-10	3.84
g) com_{coop_param} > hum_{coop_param}							
Angular gyrus	474	.001	R	42	-64	52	4.25
Angular gyrus			R	48	-64	34	4.07
Inferior frontal gyrus	201	.051	R	44	30	24	4.13

Table S4. Analysis based on the contingency of the agent's gaze behavior. The number of joint attention (JA) trials per interaction block was used as a parametric regressor. Both the contrasts referring to increasing numbers of JA as well as increasing numbers of non-joint attention (NJA) trials were reported to obtain information about the neural integration of positive and negative contingency of gaze reactions.

Region	Cluster		Side	MNI Coordinates			T
	Size	$p_{FWE\text{-}corr}$		*x*	*y*	*z*	
a) naïve$_{increase_JA}$							
Paracentral lobule	217	.002	R	4	-32	54	3.89
Paracentral lobule			L	-2	-22	58	3.63
b) naïve$_{increase_NJA}$							
Superior occipital gyrus	277	.009	R	30	-78	20	4.09
Middle occipital gyrus			R	30	-72	32	3.97
c) coop$_{increase_JA}$							
Caudate nucleus	3048	.000	L	-16	14	-6	5.38
Anterior cingulate cortex			R	2	20	22	5.15
Caudate nucleus			R	22	20	4	5.02
Putamen			L	-18	14	0	4.98
Thalamus	495	.000	R	18	-16	12	4.20
Thalamus			R	0	22	6	3.98
d) coop$_{increase_NJA}$							
Precuneus	2493	.000	R	10	-58	48	6.20
Intraparietal sulcus			R	40	-46	44	5.53
Superior parietal lobule			R	36	-60	62	5.09
Supramarginal gyrus			R	46	-36	40	4.96
Middle frontal gyrus	932	.000	R	50	24	34	5.91
Middle frontal gyrus			R	36	12	60	4.49
Intraparietal sulcus	878	.000	L	-36	-56	40	4.77
Inferior frontal gyrus	869	.000	R	32	60	8	5.10
Lateral orbital gyrus			R	44	48	-14	4.68

Table S5. Event-related analyses of gaze reactions compared JA with NJA trials.

Region	Cluster		Side	MNI Coordinates			T
	Size	$p_{FWE\text{-}corr}$		*x*	*y*	*z*	
a) $JA_{naïve}$ > $NJA_{naïve}$							
Precentral gyrus	3207	.000	R	26	-22	60	5.86
Postcentral gyrus			R	32	-30	62	5.31
Postcentral gyrus			L	-20	-40	56	5.05
Middle cingulate cortex			R	12	-16	42	4.66
Middle cingulate cortex			L	-8	-2	34	4.60
Paracentral lobule			R	4	-24	48	4.57
Paracentral lobule			L	-8	-24	48	4.56
Amygdala	510	.001	L	-22	-6	-14	4.82
Medial orbitofrontal cortex			L	-22	8	-22	4.28
Putamen			L	-20	8	-6	3.80
Thalamus	410	.003	L	-4	-18	14	4.66
Superior temporal sulcus (anterior region)	398	.004	L	-46	-6	-8	4.59
Temporal pole			L	-54	8	-14	3.93
Ventromedial prefrontal cortex	364	.006	R	4	44	-16	5.17
Anterior cingulate cortex			R	4	36	10	3.90
Anterior cingulate cortex			L	-4	38	2	3.61
Parietal operculum	249	.030	L	-40	-24	26	4.59
Parietal operculum	227	.042	R	40	-36	24	5.61
b) $NJA_{naïve}$ > $JA_{naïve}$							
Supramarginal gyrus	1768	.000	R	38	-42	42	5.31
Precuneus			R	10	-58	48	5.05
Precentral sulcus	1528	.000	L	-26	-4	58	5.68
Precentral gyrus			L	-38	4	54	5.44
Superior fontal sulcus	1458	.000	R	32	4	44	5.41
Middle frontal gyrus			R	46	24	36	5.30
Intraparietal sulcus	419	.000	L	-40	-46	36	4.40
Superior frontal gyrus	402	.000	L	-2	22	56	4.47
Inferior frontal gyrus (p. orbitalis)	220	.005	R	30	58	-6	3.97
Lateral orbital gyrus	217	.005	L	-44	44	10	4.25

Supplementary Figures

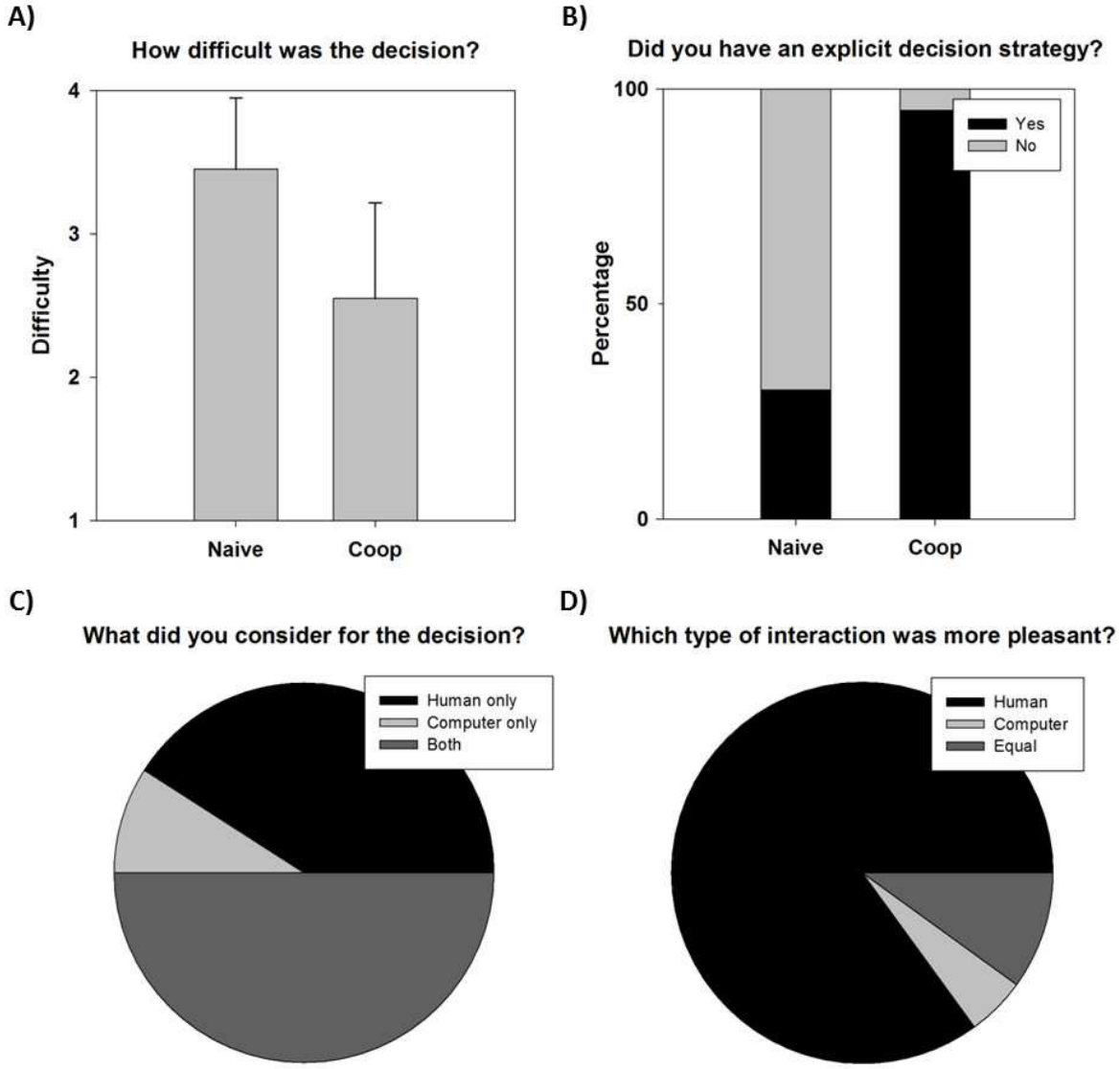

Fig. S1. Participants were asked about their experience of the task in post-experiment questionnaires. A) Results indicate that the decision was more difficult in the naïve than in the cooperative context (naïve: M = 3.45, SD = 0.49; coop: M = 2.55, SD = 0.67, $p < .001$). B) During interaction with a naïve partner, the decision was not based on strategic thinking and explicit decision criteria, but on intuitive processes. C) Participants preferred grounding their decision on considerations about human behavior rather than about the functionality of a computer program. D) Interactions that were experienced as social interactions with another human participant were experienced as more enjoyable than non-social interactions.

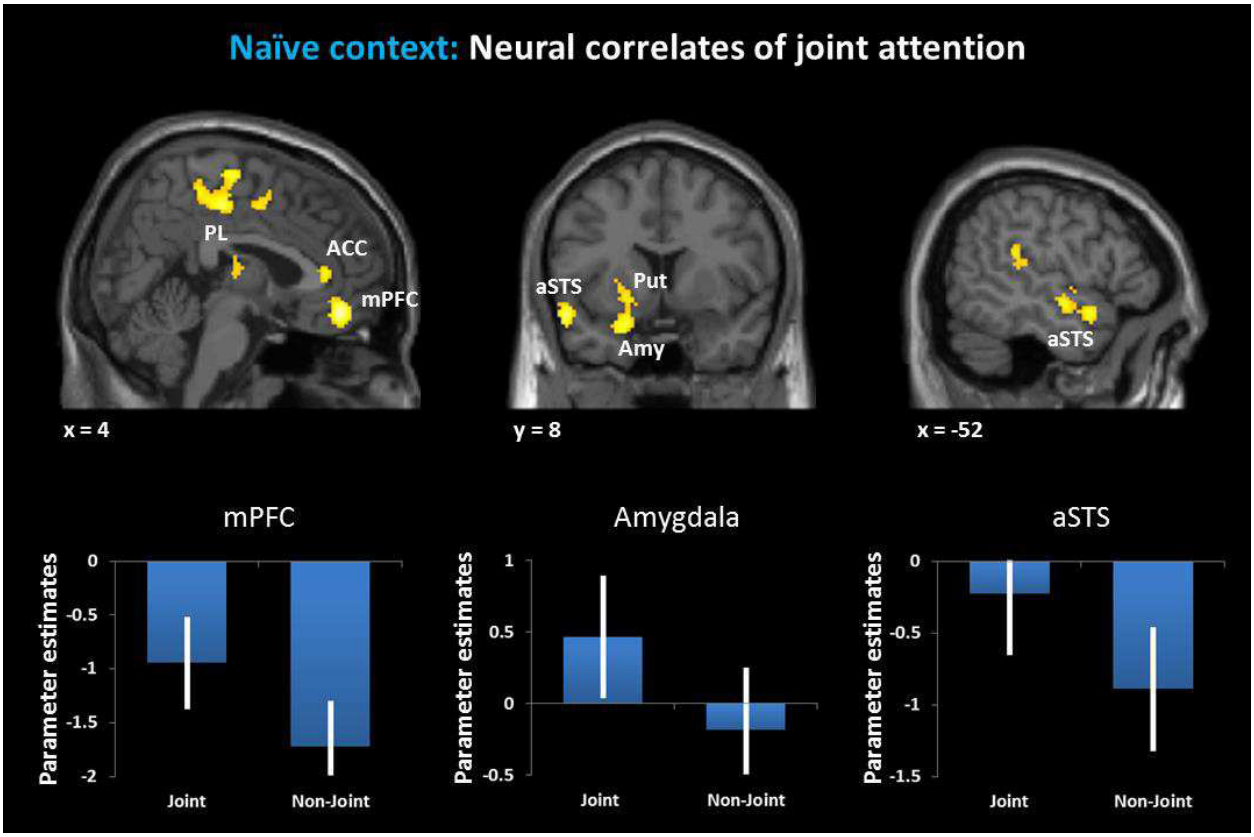

Fig. S2. An event-related analysis of joint attention trials shows activation in regions of the social brain network including the medial prefrontal cortex (mPFC), amygdala (Amy), and the anterior region of the superior temporal sulcus (aSTS). In addition, activity in the paracentral lobule (PL) was enhanced (threshold: $p < .05$ cluster-level-corrected for multiple comparisons; error bars depict 90% CI).

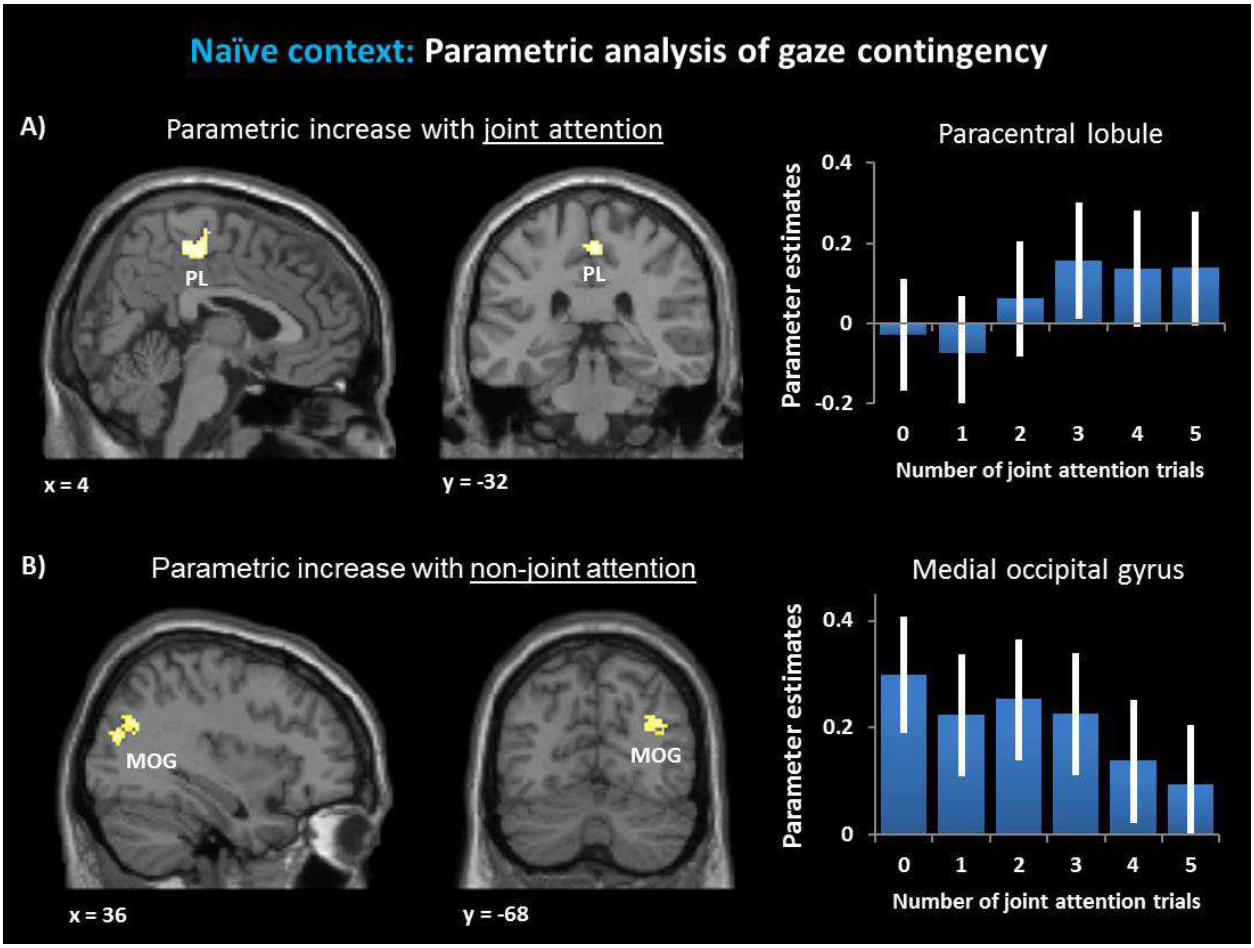

Fig. S3. Integration of gaze reactions over time in the naïve interaction context A) Parametric increases of activity with increasing numbers of joint attention trials per block in the paracentral lobule (PL). B) Parametric increase with increasing numbers of non-joint attention trials per block in the medial occipital gyrus, which corresponds to area V3 of the visual cortex (threshold: $p < .05$ cluster-level-corrected for multiple comparisons; error bars depict 90% CI).

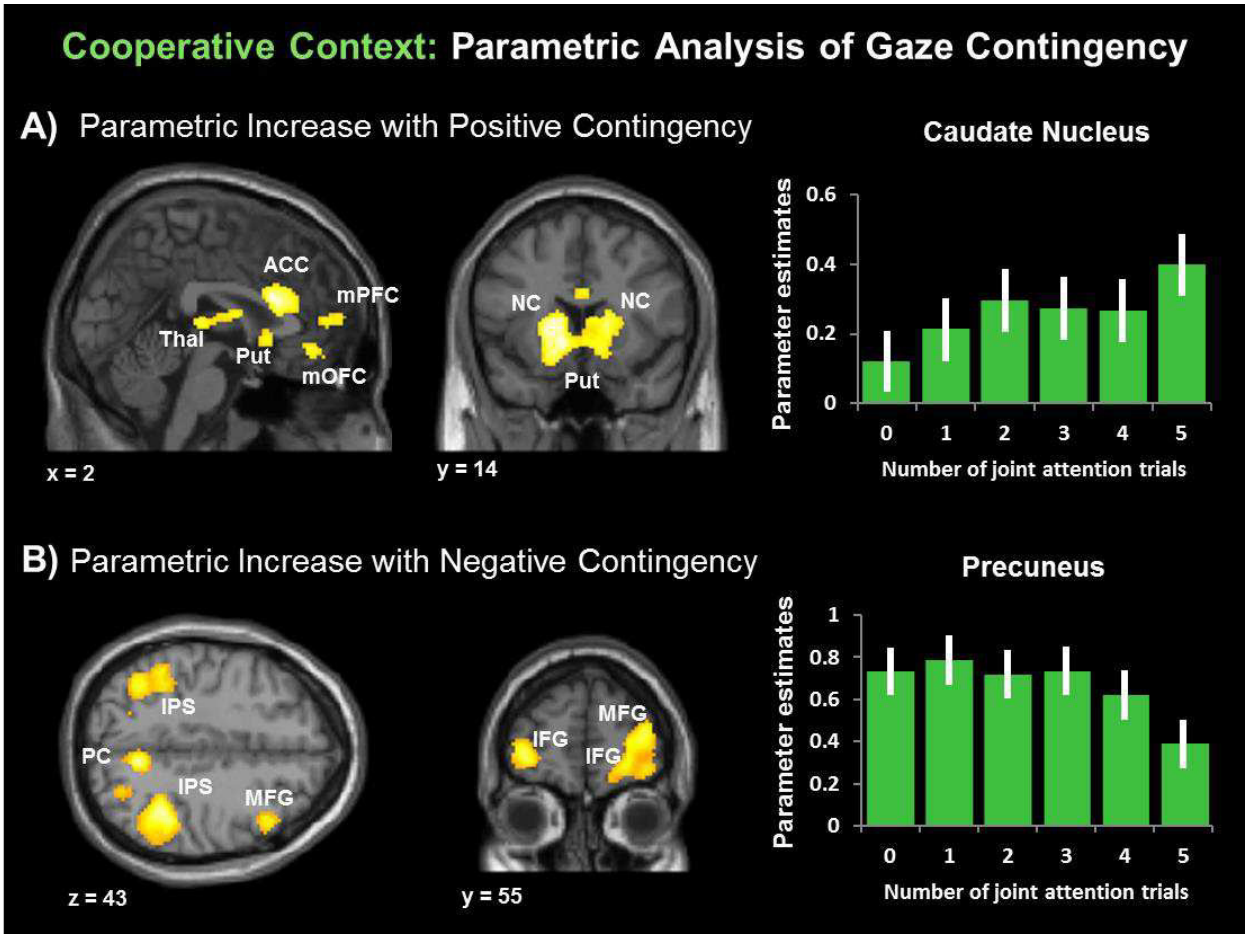

Fig. S4. Integration of gaze reactions over time during cooperative interaction. A) Parametric increases of activity with increasing numbers of joint attention trials per block in the caudate nucleus (NC), the putamen (Put), the thalamus (Thal), the anterior cingulate cortex (ACC), and the medial prefrontal cortex (mPFC). B) Parametric increases of activity with increasing numbers of non-joint attention trials per block in the FPAN including the precuneus (PC), intraparietal sulcus (IPS), and regions of the IFG and MFG corresponding to the dorso- and ventrolateral prefrontal cortex. Parameter estimates are only shown for the maxima of the largest clusters (threshold: $p < .05$ cluster-level-corrected for multiple comparisons; error bars depict 90% CI).

Notes and References

30. M. Wilms et al., It's in your eyes--using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience, Soc. Cogn. Affect. Neurosci. **5**, 98–107 (2010).

31. L. Schilbach et al., Eyes on me: an fMRI study of the effects of social gaze on action control, Soc. Cogn. Affect. Neurosci. (2010).

32. E. Redcay et al., Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience, Neuroimage **50**, 1639–1647 (2010).

33. J. Ashburner, K. J. Friston, Nonlinear spatial normalization using basis functions, Human Brain Mapping **7**, 254–266 (1999).

34. D. L. Collins, P. Neelin, T. M. Peters, A. C. Evans, Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space, J. Comput. Assist. Tomogr. **18**, 192–205 (1994).

35. V. D. Calhoun, M. C. Stevens, G. D. Pearlson, K. A. Kiehl, fMRI analysis with the general linear model: removal of latency-induced amplitude bias by incorporation of hemodynamic derivative terms, Neuroimage **22**, 252–257 (2004).

36. P. M. Macey, K. E. Macey, R. Kumar, R. M. Harper, A method for removal of global effects from fMRI time series, Neuroimage **22**, 360–366 (2004).

37. K. J. Worsley et al., A unified statistical approach for determining significant signals in images of cerebral activation, Hum. Brain Mapp. **4**, 58–73 (1996).

38. S. B. Eickhoff et al., A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data, Neuroimage **25**, 1325–1335 (2005).

39. H. M. Duvernoy, The Human Brain: Surface, Three-Dimensional Sectional Anatomy with MRI, and Blood Supply (Springer, Heidelberg, ed. 2, 1999).

40. J. H. McDonald, Handbook of Biological Statistics (Sparky House Publishing, Baltimore, MD, ed. 2nd, 2009).

41. J. Fox, D. Arena, J. N. Bailenson, Virtual Reality, J. Media Psychol. **21**, 95–113 (2009).

42. K. Vogeley, G. Bente, "Artificial humans": Psychology and neuroscience perspectives on embodiment and nonverbal communication, Neural Netw. **23**, 1077–1090 (2010).

43. C. L. Hoyt, J. Blascovich, K. R. Swinth, Social Inhibition in Immersive Virtual Environments, Presence-Teleop. Virt. **12**, 183–195 (2003).

44. J. N. Bailenson, J. Blascovich, A. C. Beall, J. M. Loomis, Interpersonal Distance in Immersive Virtual Environments, Pers. Soc. Psychol. Bull. **29**, 819–833 (2003).

45. O. Gillath, C. McCall, P. R. Shaver, J. Blascovich, What Can Virtual Reality Teach Us About Prosocial Tendencies in Real and Virtual Environments?, Media Psychol. **11**, 259–282 (2008).

46. G. Bente, F. Eschenburg, N. C. Krämer, in Virtual Reality, R. Shumaker, Ed. (Springer, Berlin, Heidelberg, 2007), vol. 4563, pp. 185–194.

47. M. Carpenter, K. Nagell, M. Tomasello, Social cognition, joint attention, and communicative competence from 9 to 15 months of age, Monogr. Soc. Res. Child Dev. **63**, i–vi, 1–143 (1998).

48. P. Mundy, L. Newell, Attention, Joint Attention, and Social Cognition, Curr. Dir. Psychol. Sci. **16**, 269–274 (2007).

49. M. Tomasello, M. Carpenter, J. Call, T. Behne, H. Moll, Understanding and sharing intentions: the origins of cultural cognition, Behav. Brain Sci. **28**, 675–691 (2005).

50. H. H. Clark, Using Language (Cambridge University Press, Cambridge, UK, ed. 2, 1996).

51. M. Tomasello, M. Carpenter, Shared intentionality, Dev. Sci. **10**, 121–125 (2007).